

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 24, 2004, 09:56:46 ; Search time 45 Seconds  
(without alignments)  
2902.767 Million cell updates/sec

Title: US-09-978-318B-1  
Perfect score: 2203  
Sequence: 1 MAALRYAGLDTDSEDELP.....RTLWALSERLQERLGSQSG 414

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum Match 100%

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SPREMBL\_25:\*

1: sp\_archaea:\*

2: sp\_bacteria:\*

3: sp\_fungi:\*

4: sp\_human:\*

5: sp\_invertebrate:\*

6: sp\_mammal:\*

7: sp\_mhc:\*

8: sp\_organelle:\*

9: sp\_phase:\*

10: sp\_plant:\*

11: sp\_rodent:\*

12: sp\_virus:\*

13: sp\_vertebrate:\*

14: sp\_unclassified:\*

15: sp\_rvirus:\*

16: sp\_bacteriaph:\*

17: sp\_archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2203	100.0	414	4 Q9NZC7	Q9nzcz7 homo sapien
2	2191	99.5	414	4 Q9NRF5	Q9nrf5 homo sapien
3	2071	94.0	414	11 Q91WL8	Q91wl8 mus musculus
4	2067	93.8	414	11 Q9JLF5	Q9jlf5 mus musculus
5	1869	84.8	363	4 Q9NRF6	Q9nrf6 homo sapien
6	1785	81.0	354	11 Q9D2B3	Q9d2b3 mus musculus
7	1777	80.7	367	11 Q9D339	Q9d339 mus musculus
8	1637.5	74.3	412	13 Q803A8	Q803a8 brachydanio
9	1157	52.5	234	4 Q9BTT8	Q9btt8 homo sapien
10	1149	52.2	242	11 Q920X2	Q920x2 mus musculus
11	1023	46.4	409	5 Q9VLU5	Q9vlu5 drosophila
12	942	42.8	191	4 Q9NRK1	Q9nrk1 homo sapien
13	927	42.1	189	4 Q9NPC9	Q9npc9 homo sapien
14	750.5	34.1	213	4 Q96KM3	Q96km3 homo sapien
15	746.5	33.9	311	4 Q96RF2	Q96rf2 homo sapien
16	740	33.6	158	11 Q8C8J6	Q8c8j6 mus musculus

17	541.5	24.6	330	5	Q960C7	Q960c7 drosophila
18	539	24.5	370	5	Q9UAW2	Q9uaw2 caenorhabdi
19	538.5	24.4	329	5	Q19062	Q19062 caenorhabdi
20	538.5	24.4	331	5	Q9V4P9	Q9v4p9 drosophila
21	537.5	24.4	330	5	Q9V4Q3	Q9v4q3 drosophila
22	537	24.4	321	10	Q9LGI8	Q9lgi8 oryza sativ
23	528	24.0	319	5	Q9SQN9	Q9sqn9 caenorhabdi
24	520.5	23.6	320	10	Q8L9T6	Q8l9t6 arabidopsis
25	520.5	23.6	325	5	Q8MZG9	Q8mzg9 drosophila
26	514.5	23.4	316	10	Q8HIQ6	Q8hiq6 arabidopsis
27	510.5	23.2	300	5	Q9V4Q2	Q9v4q2 drosophila
28	509.5	23.1	296	5	Q8MKN1	Q8mkn1 drosophila
29	506.5	23.0	293	11	Q9RIR9	Q9rir9 mus musculu
30	506.5	23.0	300	11	Q9RIR8	Q9rir8 mus musculu
31	497	22.6	317	10	Q9LDY7	Q9ldy7 arabidopsis
32	497	22.6	321	10	Q80924	Q80924 arabidopsis
33	495	22.5	331	10	Q8RWJ2	Q8rwj2 arabidopsis
34	493	22.4	319	10	Q94LU0	Q94lu0 oryza sativ
35	493	22.4	319	10	Q7XCH3	Q7xch3 oryza sativ
36	490.5	22.3	370	4	Q12953	Q12953 homo sapien
37	489	22.2	331	10	Q8LCE7	Q8lce7 arabidopsis
38	485.5	22.0	309	13	Q7T348	Q7t348 brachydanio
39	485.5	22.0	406	5	Q9W404	Q9w404 drosophila
40	478	21.7	379	5	Q45680	Q45680 caenorhabdi
41	476.5	21.6	350	10	Q9LZ48	Q9lzt48 arabidopsis
42	472	21.4	336	16	Q9RR99	Q9rr99 deinococcus
43	469	21.3	314	10	Q7XMT5	Q7xmt5 oryza sativ
44	469	21.3	320	10	P81259	P81259 brassica na
45	447.5	20.3	294	10	O81739	O81739 arabidopsis

## ALIGNMENTS

RESULT 1

Q9NZC7 PRELIMINARY; PRT; 414 AA.

ID Q9NZC7; 01-OCT-2000 (TrEMBLrel. 15, Created)

AC Q9NZC7; 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)

DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)

DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)

DE WW domain-containing protein WWOX (WW domain-containing oxidoreductase isoform FORII).

DE isoform FORII).

GN WWOX.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OX NCBI\_TaxID=9606;

RN [1]\_TaxID=9606;

RP SEQUENCE FROM N.A.

RX MEDLINE=20246348; PubMed=10786676;

RA Bednarek A.K., Laflin K.J., Daniel R.L., Liao Q., Hawkins K.A.,

RA Aldaz C.M.;

RT "WWOX, a novel WW domain-containing protein mapping to human

RT chromosome 16q23.3-24.1, a region frequently affected in breast

RT cancer.";

RL Cancer Res. 60:2140-2145(2000).

RN [2]

RN SEQUENCE FROM N.A.

RP PubMed=11572989;

RX Paige A.J.W., Taylor K.J., Taylor C., Hallier S.G., Farrington S.,

RA Scott D., Porteous D.J., Smyth J.F., Gabra H., Watson J.E.V.;

RT "WWOX: A candidate tumor suppressor gene involved in multiple tumor

RT types.";

RL Proc. Natl. Acad. Sci. U.S.A. 98:11417-11422(2001).

CC - - SIMILARITY: BELONGS TO THE SHORT-CHAIN DEHYDROGENASES/REDUCTASES

CC (SDR) FAMILY.

DR EMBL; AF211943; AAF27049.1; -

DR EMBL; AF325432; AAL05449.1; -

DR EMBL; AF325432; AAL05449.1; JOINED.

DR EMBL; AF325424; AAL05449.1; JOINED.

DR EMBL; AF325425; AAL05449.1; JOINED.

DR EMBL; AF325426; AAL05449.1; JOINED.

DR EMBL; AF325427; AAL05449.1; JOINED.  
 DR EMBL; AF325428; AAL05449.1; JOINED.  
 DR EMBL; AF325430; AAL05449.1; JOINED.  
 DR EMBL; AF325431; AAL05449.1; JOINED.  
 DR HSSP; Q13526; IPIN.  
 DR GO; GO:0005489; F:electron transporter activity; TAS.  
 DR GO; GO:0005515; F:protein binding; TAS.  
 DR GO; GO:0008202; P:steroid metabolism; TAS.  
 DR InterPro; IPR002198; ADH short.  
 DR InterPro; IPR002198; ADH short.  
 DR Pfam; PF00106; adh\_short; 1.  
 DR Pfam; PF00397; WW; 2.  
 DR SMART; SM00456; WW; 2.  
 DR PROSITE; PS01159; WW DOMAIN\_1; 2.  
 DR PROSITE; PS00200; WW DOMAIN\_2; 2.  
 DR Oxidoreductase.  
 KW Oxidoreductase.  
 SQ SEQUENCE 414 AA; 46676 MW; E4D9A649B6CB05DF CRC64;

Query Match 100.0%; Score 2203; DB 4; Length 414;  
 Best Local Similarity 100.0%; Pred. No. 1.5e-176;  
 Matches 414; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAALRYAGLDLDDTSEDELPPEWEERTTKDGVVYANHTTEKTOWEHPKTKRKRKRVAGDLP 60  
 DB 1 MAALRYAGLDLDDTSEDELPPEWEERTTKDGVVYANHTTEKTOWEHPKTKRKRKRVAGDLP 60  
 QY 61 YGWEQETDENGQVFFVDHINKRTTYLDPRLAFTVDDNPTKPTRQRYDGSSTTAMEILQGR 120  
 DB 61 YGWEQETDENGQVFFVDHINKRTTYLDPRLAFTVDDNPTKPTRQRYDGSSTTAMEILQGR 120  
 QY 121 DFTGKVVVVTGANGSIGGFETAKSFALHGAHVILACRNMARASEAVSRILEEWHKAKVEAM 180  
 DB 121 DFTGKVVVVTGANGSIGGFETAKSFALHGAHVILACRNMARASEAVSRILEEWHKAKVEAM 180  
 QY 181 TLDLALLRSVOHFAEAFKAKNPLHLVLCNAATFALPWSLTKDGLTTFQVNLGHFYL 240  
 DB 181 TLDLALLRSVOHFAEAFKAKNPLHLVLCNAATFALPWSLTKDGLTTFQVNLGHFYL 240  
 QY 241 QLLQDVLCSPARVIVVSSSHRFTDINDSLGKLDPSRLSPKNDYWLAMLAYNRSKLCN 300  
 DB 241 QLLQDVLCSPARVIVVSSSHRFTDINDSLGKLDPSRLSPKNDYWLAMLAYNRSKLCN 300  
 QY 301 ILFNSNELHRLSPRGVTSNAVHPGNMYSNIHRSWWVYTLFLTPARPTKSMQCGAATTV 360  
 DB 301 ILFNSNELHRLSPRGVTSNAVHPGNMYSNIHRSWWVYTLFLTPARPTKSMQCGAATTV 360  
 QY 361 YCAAVPELEGGMVFNCCRCMPSPQAQSEETARTLWALSRLIQLERLGSQSG 414  
 DB 361 YCAAVPELEGGMVFNCCRCMPSPQAQSEETARTLWALSRLIQLERLGSQSG 414

RESULT 2  
 Q9NRF5 PRELIMINARY; PRT; 414 AA.  
 ID Q9NRF5  
 AC Q9NRF5  
 DT 01-OCT-2000 (TRENBLrel. 15, Created)  
 DT 01-OCT-2000 (TRENBLrel. 15, Last sequence update)  
 DE 01-JUN-2003 (TRENBLrel. 24, Last annotation update)  
 DE For II protein.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OC NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=20320695; PubMed=10861292;  
 RA Ried K., Finniss M., Hobson L., Mangelsdorf M., Dayan S.,  
 RA Nancarrow J.K., Woolatt E., Kremmidiotis G., Gardner A., Venter D.,  
 RA Baker E., Richards R.I.;  
 RT "Common chromosomal fragile site FRA16D sequence: Identification of the  
 RT POR gene spanning FRA16D and homozygous deletions and translocation  
 RT breakpoints in cancer cells.";  
 RT Cancer 9:1229-1233 (1998)

CC -!- SIMILARITY: BELONGS TO THE SHORT-CHAIN DEHYDROGENASES/REDUCTASES  
 CC (SDR) FAMILY.  
 CC EMBL; AF227527; AAF82054.1; --  
 DR HSSP; Q13526; IPIN.  
 DR Genew; HGNC:12799; WWOX.  
 DR GO; GO:0016491; F:oxidoreductase activity; IEA.  
 DR GO; GO:0008152; P:metabolism; IEA.  
 DR InterPro; IPR002198; ADH short.  
 DR InterPro; IPR002198; ADH short.  
 DR Pfam; PF00106; adh\_short; 1.  
 DR Pfam; PF00397; WW; 2.  
 DR SMART; SM00456; WW; 2.  
 DR PROSITE; PS01159; WW DOMAIN\_1; 2.  
 DR PROSITE; PS00200; WW DOMAIN\_2; 2.  
 DR Oxidoreductase.  
 KW Oxidoreductase.  
 SQ SEQUENCE 414 AA; 46680 MW; 04D9B358F87F0E75 CRC64;

Query Match 99.5%; Score 2191; DB 4; Length 414;  
 Best Local Similarity 99.5%; Pred. No. 1.6e-175;  
 Matches 412; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 MAALRYAGLDLDDTSEDELPPEWEERTTKDGVVYANHTTEKTOWEHPKTKRKRKRVAGDLP 60  
 DB 1 MAALRYAGLDLDDTSEDELPPEWEERTTKDGVVYANHTTEKTOWEHPKTKRKRKRVAGDLP 60  
 QY 61 YGWEQETDENGQVFFVDHINKRTTYLDPRLAFTVDDNPTKPTRQRYDGSSTTAMEILQGR 120  
 DB 61 YGWEQETDENGQVFFVDHINKRTTYLDPRLAFTVDDNPTKPTRQRYDGSSTTAMEILQGR 120  
 QY 121 DFTGKVVVVTGANGSIGGFETAKSFALHGAHVILACRNMARASEAVSRILEEWHKAKVEAM 180  
 DB 121 DFTGKVVVVTGANGSIGGFETAKSFALHGAHVILACRNMARASEAVSRILEEWHKAKVEAM 180  
 QY 181 TLDLALLRSVOHFAEAFKAKNPLHLVLCNAATFALPWSLTKDGLTTFQVNLGHFYL 240  
 DB 181 TLDLALLRSVOHFAEAFKAKNPLHLVLCNAATFALPWSLTKDGLTTFQVNLGHFYL 240  
 QY 241 QLLQDVLCSPARVIVVSSSHRFTDINDSLGKLDPSRLSPKNDYWLAMLAYNRSKLCN 300  
 DB 241 QLLQDVLCSPARVIVVSSSHRFTDINDSLGKLDPSRLSPKNDYWLAMLAYNRSKLCN 300  
 QY 301 ILFNSNELHRLSPRGVTSNAVHPGNMYSNIHRSWWVYTLFLTPARPTKSMQCGAATTV 360  
 DB 301 ILFNSNELHRLSPRGVTSNAVHPGNMYSNIHRSWWVYTLFLTPARPTKSMQCGAATTV 360  
 QY 361 YCAAVPELEGGMVFNCCRCMPSPQAQSEETARTLWALSRLIQLERLGSQSG 414  
 DB 361 YCAAVPELEGGMVFNCCRCMPSPQAQSEETARTLWALSRLIQLERLGSQSG 414

RESULT 3  
 Q91WL8 PRELIMINARY; PRT; 414 AA.  
 ID Q91WL8  
 AC Q91WL8  
 DT 01-DEC-2001 (TRENBLrel. 19, Created)  
 DT 01-DEC-2001 (TRENBLrel. 19, Last sequence update)  
 DE 01-OCT-2003 (TRENBLrel. 25, Last annotation update)  
 DE WW domain-containing oxidoreductase.  
 GN WWOX.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OC NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX TISSUE=Kidney;  
 RA Strausberg R.;  
 RN Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.  
 RP SEQUENCE FROM N.A.  
 RP STRAIN=C57BL/6J; TISSUE=Mullerian duct includes surrounding region;  
 RC MEDLINE=22354683; PubMed=12466851;  
 RX

GenCore version 5.1.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 24, 2004, 21:49:29 ; Search time 8833 Seconds  
(without alignments)  
11109.321 Million cell updates/sec

Title: US-09-978-318B-2  
Perfect score: 2264  
Sequence: 1 gcagtcgcagcgagcg.....aaaaaaaaaaaaaaaaaaaaa 2264

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : GenEmbl.\*

- 1: gb.ba.\*
- 2: gb.htg.\*
- 3: gb.in.\*
- 4: gb.om.\*
- 5: gb.ov.\*
- 6: gb.pat.\*
- 7: gb.ph.\*
- 8: gb.pl.\*
- 9: gb.pr.\*
- 10: gb.ro.\*
- 11: gb.sts.\*
- 12: gb.sy.\*
- 13: gb.un.\*
- 14: gb.vi.\*
- 15: em.ba.\*
- 16: em.fun.\*
- 17: em.hum.\*
- 18: em.in.\*
- 19: em.mu.\*
- 20: em.om.\*
- 21: em.or.\*
- 22: em.ov.\*
- 23: em.pat.\*
- 24: em.ph.\*
- 25: em.pl.\*
- 26: em.ro.\*
- 27: em.sts.\*
- 28: em.un.\*
- 29: em.vi.\*
- 30: em.htg.hum.\*
- 31: em.htg.inv.\*
- 32: em.htg.other.\*
- 33: em.htg.mus.\*
- 34: em.htg.pln.\*
- 35: em.htg.rtd.\*
- 36: em.htg.mam.\*
- 37: em.htg.vrt.\*
- 38: em.sy.\*
- 39: em.htgo.hum.\*
- 40: em.htgo.mus.\*
- 41: em.htgo.other.\*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description	
1	2264	100.0	2264	9	AF211943	AF211943 Homo sapi	
2	2253.4	99.5	2256	6	AX877089	AX877089 Sequence	
3	2253.4	99.5	2256	6	BD012189	BD012189 A novel g	
4	2253.4	99.5	2256	6	BD156478	BD156478 Primer fo	
5	2253.4	99.5	2256	6	AK027626	AK027626 Homo sapi	
6	2213.4	97.8	2219	9	AF227527	AF227527 Homo sapi	
7	1372.4	60.6	1475	9	HS013395	UI3395 Human oxido	
8	1174	51.9	1732	9	AF395123	AF395123 Homo sapi	
9	1137.4	50.2	1225	9	AF227526	AF227526 Homo sapi	
10	1096.6	48.4	2269	10	BC014716	BC014716 Mus muscu	
11	1092.8	48.3	1669	9	BC003184	BC003184 Homo sapi	
12	1084	47.9	1625	9	AF395124	AF395124 Homo sapi	
13	1077	47.6	2197	10	AF187014	AF187014 Mus muscu	
14	1062.2	46.9	168083	9	AC092376	AC092376 Homo sapi	
c	15	1052.6	46.5	161988	2	AC009129	AC009129 Homo sapi
16	691.6	30.5	768	6	AX868436	AX868436 Sequence	
17	691.6	30.5	768	6	BD148498	BD148498 Primer fo	
18	624.8	27.6	1914	5	BC044560	BC044560 Danilo rer	
c	19	617.6	27.3	74586	2	AC060793	AC060793 Homo sapi
20	616.4	27.2	711	9	AF227528	AF227528 Homo sapi	
21	567.2	25.1	729	9	AF187015	AF187015 Homo sapi	
22	518.2	22.9	705	12	BT007739	BT007739 Synthetic	
23	516.6	22.8	705	9	BT007445	BT007445 Homo sapi	
c	24	483.6	21.4	541	6	AX873488	AX873488 Sequence
c	25	483.6	21.4	541	6	BD153550	BD153550 Primer fo
26	433.6	19.2	505	9	F325423S11	AF325432 Homo sapi	
c	27	416.8	18.4	75852	2	AC021253	AC021253 Homo sapi
28	342.4	15.1	1763	3	AK115111	AK115111 Ciona inc	
29	280.6	12.4	298	11	G15863	G15863 human STS C	
30	270.4	11.9	1460	3	AY119574	AY119574 Drosophil	
31	266	11.7	457	9	F325423S09	AF325431 Homo sapi	
32	266	11.7	147101	9	AC109134	AC109134 Homo sapi	
33	266	11.7	270752	9	AF217491S3	AF217490 Homo sapi	
34	249.4	11.0	258912	2	AC129777	AC129777 Mus muscu	
c	35	237.8	10.5	254	11	G14545	G14545 human STS S
36	235.2	10.4	74586	2	AC060793	AC060793 Homo sapi	
37	232	10.2	431	9	F325423S01	AF325423 Homo sapi	
38	232	10.2	1440	9	AF212843	AF212843 Homo sapi	
c	39	232	10.2	167015	9	AC079414	AC079414 Homo sapi
c	40	232	10.2	167195	9	AC009044	AC009044 Homo sapi
c	41	211.6	9.3	220157	2	AC098432	AC098432 Rattus no
c	42	211.6	9.3	22465	2	AC130071	AC130071 Rattus no
c	43	211.6	9.3	255094	2	AC098195	AC098195 Rattus no
44	211	9.3	524	9	AF227529	AF227529 Homo sapi	
45	206.8	9.1	778	10	AY046553S3	AY046555 Mus muscu	

ALIGNMENTS

RESULT 1  
AF211943  
LOCUS AF211943 2264 bp mRNA linear PRI 05-MAY-2000  
DEFINITION Homo sapiens WW domain-containing protein WWOX mRNA, complete cds.  
ACCESSION AF211943  
VERSION AF211943.1 GI:6729682  
KEYWORDS Homo sapiens (human)  
SOURCE Homo sapiens  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 2264)  
AUTHORS Bednarek,A.K., Laflin,K.J., Daniel,R.L., Liao,Q., Hawkins,K.A. and Aldaz,C.M.  
TITLE WWOX, a novel WW domain-containing protein mapping to human

chromosome 16q23.3-24.1, a region frequently affected in breast

Cancer Res. 60 (8), 2140-2145 (2000)

20246348

10786676

2 (bases 1 to 2264)

Bednarek, A.K. and Aldaz, C.M.

Direct Submission

Submitted (06-DEC-1999) Carcinogenesis, University of Texas M.D.

Anderson Cancer Center, Science Park-Research Division, Park Road

1C POB389, Smithville, TX 78957, USA

Location/Qualifiers

1. .2264

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/chromosome="16"

/map="16q23.3-q24.1"

1. .125

/notes="putative hydroxysteroid dehydrogenase"

/codon\_start=1

/product="WW domain-containing protein WWOX"

/protein\_id="AA27049.1"

/db\_xref="GI:6729683"

/translation="MAALRYAGLDDTDSDELPPWERTTKDGVVYVYHTEKQW

EHPTGRKRVAGDLPYGMQETDENGQVFDHINKRTYLDLPRIATVDDNPTPT

TRQYDGSRTAMETLQGRDFTGVVVTGANSIGIFETAKSFALHGHVILACRNMAR

ASRVRILHEWKAKEAMTDLALRSVQHFAEAFKAKNPLHLVLCNAATFALPFW

SLTDGLETTFQVNLGHFLVQLDOLVCSAPARVIVVSSSHRFTDINDSLGLD

FSRLSPKTDYWMALNRSKLNILFSLNELHRLSPRGVTSNAVHPGNWYNIHRS

WVYVTLFLTLRPFKSMQGAATVYCAAVPELGLGGMVFNCCRCMPSPRAQSEE

TARTLWLSERLIQERLSQSG"

1371. .2264

2217. .2222

3'UTR

polya\_signal

ORIGIN

Query Match 100.0%; Score 2264; DB 9; Length 2264;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 2264; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 GCAGTGGCAGGCGTACGGGTGCGGGCCCGGCGCGCGGTCTCTGTTGGAGCGGAG 60  
 1 GCAGTGGCAGGCGTACGGGTGCGGGCCCGGCGCGCGGTCTCTGTTGGAGCGGAG 60  
 61 TGAGTTCCTGACGAGTGCACCGCGCAGCGGCGGATAGGGGGCCAGGTGCCTCCACAGT 120  
 61 TGAGTTCCTGACGAGTGCACCGCGCAGCGGCGGATAGGGGGCCAGGTGCCTCCACAGT 120  
 121 CAGCCATGGCAGCGTGCCTACGCGGGGCTGACGACGACGACGACGACGACGACGACG 180  
 121 CAGCCATGGCAGCGTGCCTACGCGGGGCTGACGACGACGACGACGACGACGACGACG 180  
 181 CTCGGGCTGGGAGGAGAACCAACAGCGGCTGGGTCTTACTACGCGCAATCAACCG 240  
 181 CTCGGGCTGGGAGGAGAACCAACAGCGGCTGGGTCTTACTACGCGCAATCAACCG 240  
 241 AGGAGAGACTCAGTGGGAACTCCAAAATCTGGAAAAGAAAACGAGTGGCAGGAGATT 300  
 241 AGGAGAGACTCAGTGGGAACTCCAAAATCTGGAAAAGAAAACGAGTGGCAGGAGATT 300  
 301 TGCCATACGATGGGAAACAGAACTGATGAGAACGACGACGACGACGACGACGACGAC 360  
 301 TGCCATACGATGGGAAACAGAACTGATGAGAACGACGACGACGACGACGACGACGAC 360  
 361 TAAATAAAGAACCACTTCTGGACCAAGACTGGCGTTTACTGTGGATGATAATCCGA 420  
 361 TAAATAAAGAACCACTTCTGGACCAAGACTGGCGTTTACTGTGGATGATAATCCGA 420  
 421 CCAAGCCAAACCAACCGGCAAGATACGACGCGCAGCACCACTGCCATGCCAATTCCTCAGG 480  
 421 CCAAGCCAAACCAACCGGCAAGATACGACGCGCAGCACCACTGCCATGCCAATTCCTCAGG 480

481 GCCGGGATTTCACTGGCAAGTGGTGTGTGCTCACTGGAGCTAAATTCAGGAATAGGGTTCC 540  
 481 GCCGGGATTTCACTGGCAAGTGGTGTGTGCTCACTGGAGCTAAATTCAGGAATAGGGTTCC 540  
 541 AAACCGCCAGTCTTTTGGCCCTCCATGGTGCAATGTGATCTTGGCCCTGCGAGCAATCG 600  
 541 AAACCGCCAGTCTTTTGGCCCTCCATGGTGCAATGTGATCTTGGCCCTGCGAGCAATCG 600  
 601 CAAAGCGGAGTGAACAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 660  
 601 CAAAGCGGAGTGAACAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 660  
 661 CAATGACCTGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 720  
 661 CAATGACCTGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 720  
 721 CCAAGAAATGCTCTTCTTCAATGTCTTGTGTGCAACGACGACCACTTGTGCTTACCCCTG 780  
 721 CCAAGAAATGCTCTTCTTCAATGTCTTGTGTGCAACGACGACCACTTGTGCTTACCCCTG 780  
 781 GTCTCAACAAAGATGCGCTGGAGACCACTTTCAGTGAATCATCTGGGGCACTTCTTACC 840  
 781 GTCTCAACAAAGATGCGCTGGAGACCACTTTCAGTGAATCATCTGGGGCACTTCTTACC 840  
 841 TTGTCCAGCTCTCCAGGATGTTTGTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 900  
 841 TTGTCCAGCTCTCCAGGATGTTTGTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 900  
 901 CCTCAGAGTCCCATCGATTTACAGATATTAACGACTCTTGGGAAAACCTGGACTTCAGTC 960  
 901 CCTCAGAGTCCCATCGATTTACAGATATTAACGACTCTTGGGAAAACCTGGACTTCAGTC 960  
 961 GCCTCTCTCCAAACAAAACGACTATTTGGGGGATGCTGGCTTATTAACAGTCCCAAGCTCT 1020  
 961 GCCTCTCTCCAAACAAAACGACTATTTGGGGGATGCTGGCTTATTAACAGTCCCAAGCTCT 1020  
 1021 GCAACATCTCTTCTCCAAACGAGTGCACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1080  
 1021 GCAACATCTCTTCTCCAAACGAGTGCACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1080  
 1081 ACAGAGTGCATCTCGGAAATATGATGATGATGATGATGATGATGATGATGATGATGATG 1140  
 1081 ACAGAGTGCATCTCGGAAATATGATGATGATGATGATGATGATGATGATGATGATGATG 1140  
 1141 CACTGCTGTTTACCTTGGCGAGGCTTTCACCAAGTCCATGCAACAGGAGCTGCACCA 1200  
 1141 CACTGCTGTTTACCTTGGCGAGGCTTTCACCAAGTCCATGCAACAGGAGCTGCACCA 1200  
 1201 CGGTGTACTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1260  
 1201 CGGTGTACTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1260  
 1261 GCTGCGGCTGCAATGCCCTCACACAGAGCTCAGAGCAAGAGACGCGCCCGGACCTGTGGG 1320  
 1261 GCTGCGGCTGCAATGCCCTCACACAGAGCTCAGAGCAAGAGACGCGCCCGGACCTGTGGG 1320  
 1321 CGCTCAGCGAGAGGTGATTCCAAGAACGCTTGGCAGCAGTCCCGCTAAGTGGAGCTCA 1380  
 1321 CGCTCAGCGAGAGGTGATTCCAAGAACGCTTGGCAGCAGTCCCGCTAAGTGGAGCTCA 1380  
 1381 GAGCGATGGGCAACACACACCGCCCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1440  
 1381 GAGCGATGGGCAACACACACCGCCCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1440  
 1441 GCCCTTCCAAATGTCCCTCCAAACAGATCCGCAAGATTAAGGAATTAAGGAGCTCA 1500  
 1441 GCCCTTCCAAATGTCCCTCCAAACAGATCCGCAAGATTAAGGAATTAAGGAGCTCA 1500  
 1501 CAACAGAGTGAATAATCTTTAAGTACATGGGAGCAGGGAATTCCTGGGTAAGTATC 1560  
 1501 CAACAGAGTGAATAATCTTTAAGTACATGGGAGCAGGGAATTCCTGGGTAAGTATC 1560  
 1561 ACTTTTCTGGGGCTGGGCTAGGCTAGGTCTCTTTTGTGTGTGTGTGTGTGTGTGTGTGTGT 1620

Db 1561 ACTTTCTGGGGCTGGGCTAGCATAGTCTCTTTGGCTTTCTGGTGGCTGTTGAA 1620  
Qy 1621 AGTAAACCTGCTGGTGTGTAGTTCCTGATCTCCCTGGAGAGACACGAAATCTC 1680  
Db 1621 AGTAAACCTGCTGGTGTGTAGTTCCTGATCTCCCTGGAGAGACACGAAATCTC 1680  
Qy 1681 TTTCTTTTACTGTTATAGATAGACCTGAGTCCCTGCTCCCATCCAGTACCAACAGG 1740  
Db 1681 TTTCTTTTACTGTTATAGATAGACCTGAGTCCCTGCTCCCATCCAGTACCAACAGG 1740  
Qy 1741 CCACCACTGACGCGGGGCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1800  
Db 1741 CCACCACTGACGCGGGGCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1800  
Qy 1801 TGCTCTCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1860  
Db 1801 TGCTCTCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1860  
Qy 1861 CAGCTTAGCAATCTGCTGGGAGACAAATCTCAGAACCTTTGTCAGCCAGTGAGTAC 1920  
Db 1861 CAGCTTAGCAATCTGCTGGGAGACAAATCTCAGAACCTTTGTCAGCCAGTGAGTAC 1920  
Qy 1921 AGTGACACCGAGGAGTAGATAGCGAGAACTACAGGTGCGAAAGTACTTGTCTAG 1980  
Db 1921 AGTGACACCGAGGAGTAGATAGCGAGAACTACAGGTGCGAAAGTACTTGTCTAG 1980  
Qy 1981 ACTCTCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2040  
Db 1981 ACTCTCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2040  
Qy 2041 TTTCTTTAGATACCTGAAAGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2100  
Db 2041 TTTCTTTAGATACCTGAAAGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2100  
Qy 2101 TTTTGGGGGCGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATG 2160  
Db 2101 TTTTGGGGGCGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATG 2160  
Qy 2161 TTTCTCTGCTTTTCAATGATGATGATGATGATGATGATGATGATGATGATGATG 2220  
Db 2161 TTTCTCTGCTTTTCAATGATGATGATGATGATGATGATGATGATGATGATGATG 2220  
Qy 2221 AAAGAACTGCTTGAATATCAAAAAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2264  
Db 2221 AAAGAACTGCTTGAATATCAAAAAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2264

RESULT 2  
AX877089 2256 bp DNA linear PAT 17-DEC-2003  
LOCUS  
DEFINITION Sequence 11994 from Patent EP1074617.  
ACCESSION AX877089  
VERSION AX877089.1 GI:40031825  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE  
1 Ota, T., Isogai, T., Nishikawa, T., Hayashi, K., Saito, K., Yamamoto, J.,  
Ishii, S., Sugiyama, T., Wakamatsu, A., Negai, K. and Otsuki, T.  
Primers for synthesizing full-length cDNA and their use  
Patent: EP 1074617-A 11994 07-FEB-2001;  
Research Association for Biotechnology (JP)  
FEATURES  
source  
1..2256  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"  
125..1369  
/note="unnamed protein product"  
/codon\_start=1  
CDS

/protein\_id="CAE89566.1"  
/db\_xref="GI:40031826"  
/translation="MAALVAGLDDTDESELPFGWEERTTKDQWVYVYANHTERTKTON  
EHPHTGKRVAGDLPGWEOTDENGQVDFVHDKRTVLDPLRAHVLACRNMAR  
TRQYDGSSTIMEILQDPDFGKVVVTGANSIGFETAKSFALHGAHVILACRNMAR  
ASEAVSRILEEMHAKAVEAMTLDLALLRSVQHFAEPAKKNVPLHVLVKNRATFALPW  
SLTDGLETTTFQVNLGHFYLQDLVLCCKSFARVIVVSSSEHFRFTDINDSKLSD  
FSRLSPDKNDYWLAVNRSLCNLFSENLHRLSPRGVTSNAPHPNMMYSNIHRS  
WVVVTLFLPTLARPTKSMQOGAAATTVCAVPELEGLGMYFNCCRCMPSPQAQSEE  
TARTLWALSERLIQERLGSQSG"  
ORIGIN  
Query Match 99.5%; Score 2253.4; DB 6; Length 2256;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 2254; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
Qy 2 CAGTGGCAGGCTGAGCGGTGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 61  
Db 1 CAGTGGCAGGCTGAGCGGTGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 60  
Qy 62 GAGTTCTCAGCGAGTGGACCCGCGCAGCGGCGGCGGCGGCGGCGGCGGCGGCGG 121  
Db 61 GAGTTCTCAGCGAGTGGACCCGCGCAGCGGCGGCGGCGGCGGCGGCGGCGGCGG 120  
Qy 122 AGCCATGGCAGCGCTGCGCTACGCGGCGGCTGCGACGACGACGACGACGACGACG 181  
Db 121 AGCCATGGCAGCGCTGCGCTACGCGGCGGCTGCGACGACGACGACGACGACGACG 180  
Qy 182 TCGGCGCTGGGAGGAGAGAACCCACGACGCGGCGGCGGCGGCGGCGGCGGCGGCGG 241  
Db 181 TCGGCGCTGGGAGGAGAGAACCCACGACGCGGCGGCGGCGGCGGCGGCGGCGGCGG 240  
Qy 242 CGAGAGAGCTCAGTGGGAGACATCCAAAACTCGAAAAAGAAAAACGAGTGGCAGGAGATT 301  
Db 241 CGAGAGAGCTCAGTGGGAGACATCCAAAACTCGAAAAAGAAAAACGAGTGGCAGGAGATT 300  
Qy 302 GCCATACCGATGGGAGAACAAAGAACTGATGAGAACCGAACGAGTGGTGGTGGTGGT 361  
Db 301 GCCATACCGATGGGAGAACAAAGAACTGATGAGAACCGAACGAGTGGTGGTGGTGGT 360  
Qy 362 AAATAAAGAACCACTTACCTTGGACCCAGACCTGGGCGTGGTGGTGGTGGTGGTGGT 421  
Db 361 AAATAAAGAACCACTTACCTTGGACCCAGACCTGGGCGTGGTGGTGGTGGTGGTGGT 420  
Qy 422 CAAAGCCAAACCCCGGCAAGATACGACGCGCAGCACCATGCGCATGCAATTCCTCAGGG 481  
Db 421 CAAAGCCAAACCCCGGCAAGATACGACGCGCAGCACCATGCGCATGCAATTCCTCAGGG 480  
Qy 482 CCGGGATTTTCACTGGCAAAAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 541  
Db 481 CCGGGATTTTCACTGGCAAAAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 540  
Qy 542 AACCGCAAGTCTTTGGCCCTCATGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 601  
Db 541 AACCGCAAGTCTTTGGCCCTCATGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 600  
Qy 602 AAGCGGAGTGAAGCAGTGTCAAGCATTTTGAAGAAATGCGATAAAGCAAGGTAGAGC 661  
Db 601 AAGCGGAGTGAAGCAGTGTCAAGCATTTTGAAGAAATGCGATAAAGCAAGGTAGAGC 660  
Qy 662 AATGACCTGGACCTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 721  
Db 661 AATGACCTGGACCTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 720  
Qy 722 CAGAATGTGCTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 781  
Db 721 CAGAATGTGCTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 780  
Qy 782 TCTCACCAGAGATGGCTGGAGACCATCTTCAAGTGAATCATCTGGGGCACTTCTACCT 841  
Db 781 TCTCACCAGAGATGGCTGGAGACCATCTTCAAGTGAATCATCTGGGGCACTTCTACCT 840  
Qy 842 TGTCCAGCTCTCCAGGATGTTTTGTGGCGGCTCAGCTCTGCGCGGTGTCATGTTGGTCTC 901

# RESULT 4

AB84485  
ID AAB84485 standard; protein; 414 AA.  
XX AC AAB84485;

05-SEP-2001 (first entry)

Amino acid sequence of FRA16D oxidoreductase (FOR) II gene cDNA.

Cancer associated protein; FOR gene; FRA16D; fragile site; aphidicolin;  
chromosomal rearrangement; cancer; splice variant; DNA instability;  
FRA16D oxidoreductase; neoplasia.

Homo sapiens.

WO200144466-A1.

21-JUN-2001.

15-DEC-2000; 2000WO-AU001539.

16-DEC-1999; 99AU-00004711.

19-APR-2000; 2000AU-00007025.

(WOME-) WOMEN'S & CHILDREN'S HOSPITAL.

Richards R, Ried K, Finnis M, Hobson L, Mangelsdorf M, Dayan S;

Nancarrow J, Woollatt E, Baker E;

WPI; 2001-398151/42.

N-PSDB; AAH27867.

Novel isolated 16q23.2 nucleic acid molecule, FRA16D oxidoreductase (FOR)  
gene associated with FRA16D site, useful for early diagnosis and  
assessment of risk of cancers associated with the FRA16D region.

Claim 31; Fig 9; 150pp; English.

The present sequence represents a human FRA16D oxidoreductase (FOR) II  
transcript. The FOR gene encodes a cancer associated protein. The FRA16D  
site is a fragile site induced by aphidicolin, which is located within  
the FOR gene. The fragile site is the location of breakpoints of a  
variety of chromosomal rearrangements, and other mutations associated  
with cancers. The FOR protein is expressed as a number of splice  
variants. FOR gene polynucleotide fragments are capable of acting as  
specific primers or probes for detecting cancer associated variations of  
DNA sequence such as a point mutation or small DNA rearrangement

Tue May 25 10:19:07 2004

us-09-

CC associated with the tumour, a breakpoint of one or more chromosomal  
rearrangements associated with the tumour and a pause site within the  
FRA16 gene. FOR nucleic acid molecules are useful as markers to identify  
CC relationship between the fragile site (FRA16D) and the DNA instability in  
CC neoplasia which allows better diagnosis of cancers associated with the  
CC region  
XX  
SQ Sequence 414 AA;

Query Match 99.5%; Score 2191; DB 4; Length 414;

Best Local Similarity 99.5%; Pred. No. 1.7e-214;

Matches 412; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 MAALRYAGLDDTSDSELPFGWEERTTKDGGVYVYANHTTEKTQWEHPKTKGKRVAGDLP 60

Db 1 MAALRYAGLDDTSDSELPFGWEERTTKDGGVYVYANHTTEKTQWEHPKTKGKRVAGDLP 60

QY 61 YGWEQETDENGQVFFVDHINKRTTYLDPLAFTVDDNPKPTTRORYDGSSTTAMEILOQR 120

Db 61 YGWEQETDENGQVFFVDHINKRTTYLDPLAFTVDDNPKPTTRORYDGSSTTAMEILOQR 120

QY 121 DFTGKVVVVTGANSIGIGFETAKSFALHGAHVILACRNMARASEAVSRILLEWHKAKVEM 180

Db 121 DFTGKVVVVTGANSIGIGFETAKSFALHGAHVILACRNMARASEAVSRILLEWHKAKVEM 180

QY 181 TLDLALLRSVQHHPARAFKANNVPLHVLVCNAATFALFWSLTQDGLTTQVNHGHFYL 240

Db 181 TLDLALLRSVQHHPARAFKANNVPLHVLVCNAATFALFWSLTQDGLTTQVNHGHFYL 240

QY 241 QLLQDVLCRSAPARVIVVSSSESRFTDINDSLGKLPFSLSPKNDYMWAMLYNRSKLCN 300

Db 241 QLLQDVLCRSAPARVIVVSSSESRFTDINDSLGKLPFSLSPKNDYMWAMLYNRSKLCN 300

QY 301 ILFSNELHRLRLSPRGVTSNAVHPGNMYSNIHRSWVYVYLLFTLARPFTKSMQCGAATV 360

Db 301 ILFSNELHRLRLSPRGVTSNAVHPGNMYSNIHRSWVYVYLLFTLARPFTKSMQCGAATV 360

QY 361 YCAAVPELEGCGMYFNCCRCMPSPQAQSEETARTLWALSERLIOERLQSQSG 414

Db 361 YCAAVPELEGCGMYFNCCRCMPSPQAQSEETARTLWALSERLIOERLQSQSG 414

# ALIGNMENTS

RESULT 1  
 ABG96559  
 ID ABG96559 standard; protein; 414 AA.  
 XX AC ABG96559;  
 XX DT 12-DEC-2002 (first entry)  
 XX DE Human short chain dehydrogenase family member WWOX.  
 XX KW Short chain dehydrogenase; SDR; human; antimicrotoca; pesticide;  
 KW herbicide; DHPF deficiency; phenylketonuria; galactosaemia III;  
 KW dienoyl CoA reductase deficiency; adrenal hyperplasia; ovarian cancer;  
 KW adrenogenital syndrome; mineralcorticoid excess syndrome; breast cancer;  
 KW male pseudohermaphroditism; Zellweger syndrome; Down's syndrome;  
 KW polycystic kidney disease; Alzheimer's disease; retinitis pigmentosa;  
 KW retinitis punctata albesens; arterial hypertension; follicular lymphoma;  
 KW hepatocarcinogenesis; fungicide; antibiotic.  
 XX OS Homo sapiens.  
 XX PN WO200212544-A2.  
 XX PD 14-FEB-2002.  
 XX PF 07-AUG-2001; 2001WO-EP009140.  
 XX PR 07-AUG-2000; 2000US-0223436P.  
 XX PA (BION-) BIONETWORKS GMBH.  
 XX PI Wilckens T;  
 XX PS WPI; 2002-241770/29.  
 XX PT Identifying or verifying members of the short chain dehydrogenase (SDR)  
 PT family, useful for novel drug development (e.g. for the development of  
 PT antimicrotoca, pesticides or herbicides), by employing an algorithm using  
 PT core SDR motifs.  
 XX PS Disclosure; Fig 4; 168pp; English.  
 XX CC The invention relates to identifying or verifying members of the short  
 CC chain dehydrogenase (SDR) family comprising employing an algorithm using  
 CC core SDR motifs (MT1-MT4 and MV1, MV2 given in the specification) for  
 CC searching members of the SDR family. Also included are a member of the  
 CC SDR family identified with the method above, a method for providing  
 CC modulators for members of the SDR family, a method for evaluation of lead

Tue May 25 10:19:07 2004

us-09-97

CC -candidates for possible modulators of a member of the SDR family and a  
 CC method for detecting clinically relevant polymorphisms or single  
 CC nucleotide polymorphisms. The method is useful for screening SDR  
 CC sequences and modulators of the SDR family. The method is especially  
 CC useful as a platform for novel drug development. The SDRs can serve for  
 CC the development of e.g. antimicrotoca, pesticides or herbicides. The  
 CC modulators may be especially useful for the prophylaxis, treatment or/and  
 CC diagnosis of diseases (e.g. DHPF deficiency, phenylketonuria, dienoyl CoA  
 CC reductase deficiency, galactosaemia III, adrenal hyperplasia,  
 CC adrenogenital syndrome, mineralcorticoid excess syndrome, ovarian cancer,  
 CC breast cancer, male pseudohermaphroditism, Zellweger syndrome, polycystic  
 CC kidney disease, Alzheimer's disease, retinitis punctata albesens,  
 CC retinitis pigmentosa, Down's syndrome, arterial hypertension, follicular  
 CC lymphoma and hepatocarcinogenesis) particularly as a fungicide or  
 CC antibiotic. The present sequence is one of 39 human SDR family members  
 CC identified by the method of the invention  
 XX SQ Sequence 414 AA;  
 Query Match 100.0%; Score 2203; DB 5; Length 414;  
 Best Local Similarity 100.0%; Pred. No. 1e-215;  
 Matches 414; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 MAALRYAGLDLDDTSDDELPPGWEERTTKDGMVYVYANHTTEKTOWEHPKTKRVRAGDLP 60  
 DB 1 MAALRYAGLDLDDTSDDELPPGWEERTTKDGMVYVYANHTTEKTOWEHPKTKRVRAGDLP 60  
 QY 61 YGWEQSTDENGQVFFVDHINKETTYLDPLAPTVDDNPTKPTTRQRYDGSITAMEILQGR 120  
 DB 61 YGWEQSTDENGQVFFVDHINKETTYLDPLAPTVDDNPTKPTTRQRYDGSITAMEILQGR 120  
 QY 121 DFTGKVVVVTGANSIGIGFETAKSFALHGAHVILACENMARASEAVSRILLEWHKAKVEM 180  
 DB 121 DFTGKVVVVTGANSIGIGFETAKSFALHGAHVILACENMARASEAVSRILLEWHKAKVEM 180  
 QY 181 TLDDLALLRSVQHPAEAFKAKVPLHLVLCNAATFALPWSLTGDLGLETTPQVNLGHFYLIV 240  
 DB 181 TLDDLALLRSVQHPAEAFKAKVPLHLVLCNAATFALPWSLTGDLGLETTPQVNLGHFYLIV 240  
 QY 241 QLQDVLCHRSAPARVTVSSSEHRTFDINDSLGKLDPSRLSPKNDYMWALNYSKLCN 300  
 DB 241 QLQDVLCHRSAPARVTVSSSEHRTFDINDSLGKLDPSRLSPKNDYMWALNYSKLCN 300  
 QY 301 ILFNSLHRLSPRGVTSNAVHPGNMYSNIHRSWMVYTLFLTPARPTKSMQCGAATTV 360  
 DB 301 ILFNSLHRLSPRGVTSNAVHPGNMYSNIHRSWMVYTLFLTPARPTKSMQCGAATTV 360  
 QY 361 YCAAVPELEGICGMYPNNCCRCMPSPQAQSEHETARTLWALSERLIQERLGSQSG 414  
 DB 361 YCAAVPELEGICGMYPNNCCRCMPSPQAQSEHETARTLWALSERLIQERLGSQSG 414



GenCore version 5.1.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 24, 2004, 09:56:46 ; Search time 45 Seconds  
(without alignments)  
2902.767 Million cell updates/sec

Title: US-09-978-318B-1

Perfect score: 2203

Sequence: 1 MAALRYAGLDDTSEDELP.....RTLWALSERLQERIGSQSG 414

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL 25.\*  
1: sp\_archaea.\*  
2: sp\_bacteria.\*  
3: sp\_fungi.\*  
4: sp\_human.\*  
5: sp\_invertebrate.\*  
6: sp\_mammal.\*  
7: sp\_mhc.\*  
8: sp\_organelle.\*  
9: sp\_phase.\*  
10: sp\_plant.\*  
11: sp\_rdent.\*  
12: sp\_virus.\*  
13: sp\_vertebrate.\*  
14: sp\_unclassified.\*  
15: sp\_virus.\*  
16: sp\_bacteriap.\*  
17: sp\_archaeap.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2203	100.0	414	4 Q9NZC7	Q9nc7 homo sapien
2	2191	99.5	414	4 Q9NRP5	Q9nrf5 homo sapien
3	2071	94.0	414	11 Q91WL8	Q91wl8 mus musculu
4	2067	93.8	414	11 Q9JLF5	Q9jlf5 mus musculu
5	1869	84.8	363	4 Q9NRF6	Q9nrf6 homo sapien
6	1785	81.0	354	11 Q9D2B3	Q9d2b3 mus musculu
7	1777	80.7	367	11 Q9D339	Q9d339 mus musculu
8	1637.5	74.3	412	13 Q803A8	Q803a8 brachydanio
9	1157	52.5	234	4 Q9BTF8	Q9btf8 homo sapien
10	1149	52.2	242	11 Q920F2	Q920f2 mus musculu
11	1023	46.4	409	5 Q9VL05	Q9vlu5 drosophila
12	942	42.8	191	4 Q9NRK1	Q9nrk1 homo sapien
13	927	42.1	189	4 Q9NPC9	Q9npc9 homo sapien
14	750.5	34.1	213	4 Q96KM3	Q96km3 homo sapien
15	746.5	33.9	311	4 Q96RP2	Q96rf2 homo sapien
16	740	33.6	158	11 Q8C8J6	Q8c8j6 mus musculu

17	541.5	24.6	330	5 Q960C7	Q960c7 drosophila
18	539	24.5	370	5 Q9UAW2	Q9uaw2 caenorhabdi
19	538.5	24.4	329	5 Q19062	Q19062 caenorhabdi
20	538.5	24.4	331	5 Q9V4P9	Q9v4p9 drosophila
21	537.5	24.4	330	5 Q9V4Q3	Q9v4q3 drosophila
22	537	24.4	321	10 Q9LGI8	Q9lg8 oryza sativ
23	528	24.0	319	5 Q95QN9	Q95qn9 caenorhabdi
24	520.5	23.6	320	10 Q8L9T6	Q8l9t6 arabidopsis
25	520.5	23.6	325	5 Q8MZG9	Q8mzg9 drosophila
26	514.5	23.4	316	10 Q8H1Q6	Q8h1q6 arabidopsis
27	510.5	23.2	300	5 Q9V4Q2	Q9v4q2 drosophila
28	509.5	23.1	296	5 Q8MKN1	Q8mkn1 drosophila
29	506.5	23.0	293	11 Q9R1R9	Q9r1r9 mus musculu
30	506.5	23.0	300	11 Q9R1R8	Q9r1r8 mus musculu
31	497	22.6	317	10 Q9LDY7	Q9ldy7 arabidopsis
32	497	22.6	321	10 Q809Z4	Q809z4 arabidopsis
33	495	22.5	331	10 Q8RWJ2	Q8rwj2 arabidopsis
34	493	22.4	319	10 Q94LU0	Q94lu0 oryza sativ
35	493	22.4	319	10 Q7XCH3	Q7xch3 oryza sativ
36	490.5	22.3	370	4 Q12953	Q12953 homo sapien
37	489	22.2	331	10 Q8LCE7	Q8lce7 arabidopsis
38	485.5	22.0	309	13 Q7T3A8	Q7t3a8 brachydanio
39	485.5	22.0	406	5 Q9W404	Q9w404 drosophila
40	478	21.7	379	5 Q45680	Q45680 caenorhabdi
41	476.5	21.6	350	10 Q9LZ48	Q9lza8 arabidopsis
42	472	21.4	336	16 Q9RR99	Q9rr99 deinococcus
43	469	21.3	314	10 Q7XMT5	Q7xmt5 oryza sativ
44	469	21.3	320	10 P81259	P81259 brassica na
45	447.5	20.3	294	10 Q81739	Q81739 arabidopsis

#### ALIGNMENTS

#### RESULT 1

Q9NZC7	PRELIMINARY;	PRT;	414 AA.
ID	Q9NZC7		
AC	Q9NZC7;		
DC	01-OCT-2000 (TREMELrel. 15, Created)		
DT	01-OCT-2000 (TREMELrel. 15, Last sequence update)		
DT	01-OCT-2003 (TREMELrel. 25, Last annotation update)		
DE	WW domain-containing protein WMOX (WW domain-containing oxidoreductase isoform FORII).		
DE	isoform FORII).		
GN	WMOX.		
OS	Homo sapiens (Human).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
OX	NCBI_taxID=9606;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RP	MEDLINE=20246348; PubMed=10786676;		
RA	Bednarek A.K., Laflin K.J., Daniel R.L., Liao Q., Hawkins K.A.,		
RA	Aldaz C.M.;		
RT	"WMOX, a novel WW domain-containing protein mapping to human		
RT	chromosome 16q23.3-24.1, a region frequently affected in breast		
RT	cancer.";		
RL	Cancer Res. 60:2140-2145(2000).		
RN	[2]		
RP	SEQUENCE FROM N.A.		
RP	PubMed=11572989;		
RA	Paige A.J.W., Taylor K.J., Taylor C., Hallier S.G., Farrington S.,		
RA	Scott D., Porteous D.J., Smyth J.F., Gabra H., Watson J.E.V.;		
RT	"WMOX: A candidate tumor suppressor gene involved in multiple tumor		
RT	types";		
RL	Proc. Natl. Acad. Sci. U.S.A. 98:11417-11422(2001).		
CC	-!- SIMILARITY: BELONGS TO THE SHORT-CHAIN DEHYDROGENASES/REDUCTASES		
CC	(SDF) FAMILY.		
DR	EMBL; AF211943; AAF27049.1; -		
DR	EMBL; AF325432; AAL05449.1; -		
DR	EMBL; AF325433; AAL05449.1; JOINED.		
DR	EMBL; AF325434; AAL05449.1; JOINED.		
DR	EMBL; AF325435; AAL05449.1; JOINED.		
DR	EMBL; AF325436; AAL05449.1; JOINED.		



DR EMBL; AF325427; AAL05449.1; JOINED.  
 DR EMBL; AF325428; AAL05449.1; JOINED.  
 DR EMBL; AF325430; AAL05449.1; JOINED.  
 DR EMBL; AF325431; AAL05449.1; JOINED.  
 DR HSSP; Q13526; 1PIN.  
 DR GO; GO:0005489; F:electron transporter activity; TAS.  
 DR GO; GO:0005515; F:protein binding; TAS.  
 DR GO; GO:0008202; P:steroid metabolism; TAS.  
 DR InterPro; IPR002198; ADH short.  
 DR InterPro; IPR001202; WW\_Rsp5\_WWP.  
 DR Pfam; PF00106; adh\_short; 1.  
 DR Pfam; PF00397; WW\_2.  
 DR SMART; SM00456; WW; 2.  
 DR PROSITE; PS01159; WW\_DOMAIN\_1; 2.  
 DR PROSITE; PS0020; WW\_DOMAIN\_2; 2.  
 KW Oxidoreductase.  
 SQ SEQUENCE 414 AA; 46676 MW; E4D9A649E6CB05DF CRC64;

Query Match 100.0%; Score 2203; DB 4; Length 414;  
 Best Local Similarity 100.0%; Pred. No. 1.5e-176;  
 Matches 414; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAALRYAGLDGDDTSEDELPPEWERTTKDGVVYANHTEKTOWEHPKTKRKRKRVAGDLP 60  
 DB 1 MAALRYAGLDGDDTSEDELPPEWERTTKDGVVYANHTEKTOWEHPKTKRKRKRVAGDLP 60

QY 61 YGWEQETDENGQVFFVDHINKRTTYLDPLAFTVDNDNPKPTTRQRYDGSSTTAMEILOQR 120  
 DB 61 YGWEQETDENGQVFFVDHINKRTTYLDPLAFTVDNDNPKPTTRQRYDGSSTTAMEILOQR 120

QY 121 DFTGKVVVVTGANGSIGFETAKSFALHGAHVILACRNMARASAVSRILEEWHKAKVEAM 180  
 DB 121 DFTGKVVVVTGANGSIGFETAKSFALHGAHVILACRNMARASAVSRILEEWHKAKVEAM 180

QY 181 TLDLALLRSVQHFAFAKAKNVPLHVLVCNAATFALPWSLTQDGLTTTQVNHHLGHFYLV 240  
 DB 181 TLDLALLRSVQHFAFAKAKNVPLHVLVCNAATFALPWSLTQDGLTTTQVNHHLGHFYLV 240

QY 241 QLLQDVLCRSAPARVIVVSESHRFTDINDSLGKLDFSRLSPKNDYMWALAYNRSKLCN 300  
 DB 241 QLLQDVLCRSAPARVIVVSESHRFTDINDSLGKLDFSRLSPKNDYMWALAYNRSKLCN 300

QY 301 ILFNSLHRLSPRGVTSNAVHPGNMYSNIHRSWWVYLLFTLAPPTKSMQOQAATTV 360  
 DB 301 ILFNSLHRLSPRGVTSNAVHPGNMYSNIHRSWWVYLLFTLAPPTKSMQOQAATTV 360

QY 361 YCAAPELEGLGGMVFNCCRCMPSEAPQSEETARTLWALSRLIQLERLGSQSG 414  
 DB 361 YCAAPELEGLGGMVFNCCRCMPSEAPQSEETARTLWALSRLIQLERLGSQSG 414

RESULT 2  
 Q9NRP5 PRELIMINARY; PRT; 414 AA.

ID Q9NRP5  
 AC Q9NRP5  
 DT 01-OCT-2000 (Tremblrel. 15, Created)  
 DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)  
 DT 01-JUN-2003 (Tremblrel. 24, Last annotation update)  
 DE For II protein.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=20320695; PubMed=10861292;  
 RA Ried K., Finnis M., Hobson L., Mangelndorf M., Dayan S.,  
 RA Nancarrow J.K., Woolatt E., Kremmidiotis G., Gardner A., Venter D.,  
 RA Baker E., Richards R.I.;  
 RT "Comon chromosomal fragile site FRA16D sequence: Identification of the  
 RT FOR gene spanning FRA16D and homozygous deletions and translocation  
 RT breakpoints in cancer cells.";  
 RL Hum. Mol. Genet. 9:1651-1663 (2000).

CC -!- SIMILARITY: BELONGS TO THE SHORT-CHAIN DEHYDROGENASES/REDUCTASES  
 (SDR) FAMILY.  
 CC EMBL; AF227527; AAF82054.1; -.  
 DR HSSP; Q13526; 1PIN.  
 DR Genew; HGNC:12799; WWOX.  
 DR GO; GO:0016491; F:oxidoreductase activity; IEA.  
 DR GO; GO:0008152; P:metabolism; IEA.  
 DR InterPro; IPR002198; ADH short.  
 DR InterPro; IPR001202; WW\_Rsp5\_WWP.  
 DR Pfam; PF00106; adh\_short; 1.  
 DR Pfam; PF00397; WW\_2.  
 DR SMART; SM00456; WW; 2.  
 DR PROSITE; PS01159; WW\_DOMAIN\_1; 2.  
 DR PROSITE; PS0020; WW\_DOMAIN\_2; 2.  
 KW Oxidoreductase.  
 SQ SEQUENCE 414 AA; 46680 MW; 04D9B358F87F0E75 CRC64;

Query Match 99.5%; Score 2191; DB 4; Length 414;  
 Best Local Similarity 99.5%; Pred. No. 1.6e-175;  
 Matches 412; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 MAALRYAGLDGDDTSEDELPPEWERTTKDGVVYANHTEKTOWEHPKTKRKRKRVAGDLP 60  
 DB 1 MAALRYAGLDGDDTSEDELPPEWERTTKDGVVYANHTEKTOWEHPKTKRKRKRVAGDLP 60

QY 61 YGWEQETDENGQVFFVDHINKRTTYLDPLAFTVDNDNPKPTTRQRYDGSSTTAMEILOQR 120  
 DB 61 YGWEQETDENGQVFFVDHINKRTTYLDPLAFTVDNDNPKPTTRQRYDGSSTTAMEILOQR 120

QY 121 DFTGKVVVVTGANGSIGFETAKSFALHGAHVILACRNMARASAVSRILEEWHKAKVEAM 180  
 DB 121 DFTGKVVVVTGANGSIGFETAKSFALHGAHVILACRNMARASAVSRILEEWHKAKVEAM 180

QY 181 TLDLALLRSVQHFAFAKAKNVPLHVLVCNAATFALPWSLTQDGLTTTQVNHHLGHFYLV 240  
 DB 181 TLDLALLRSVQHFAFAKAKNVPLHVLVCNAATFALPWSLTQDGLTTTQVNHHLGHFYLV 240

QY 241 QLLQDVLCRSAPARVIVVSESHRFTDINDSLGKLDFSRLSPKNDYMWALAYNRSKLCN 300  
 DB 241 QLLQDVLCRSAPARVIVVSESHRFTDINDSLGKLDFSRLSPKNDYMWALAYNRSKLCN 300

QY 301 ILFNSLHRLSPRGVTSNAVHPGNMYSNIHRSWWVYLLFTLAPPTKSMQOQAATTV 360  
 DB 301 ILFNSLHRLSPRGVTSNAVHPGNMYSNIHRSWWVYLLFTLAPPTKSMQOQAATTV 360

QY 361 YCAAPELEGLGGMVFNCCRCMPSEAPQSEETARTLWALSRLIQLERLGSQSG 414  
 DB 361 YCAAPELEGLGGMVFNCCRCMPSEAPQSEETARTLWALSRLIQLERLGSQSG 414

RESULT 3  
 Q91WL8 PRELIMINARY; PRT; 414 AA.

ID Q91WL8  
 AC Q91WL8  
 DT 01-DEC-2001 (Tremblrel. 19, Created)  
 DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)  
 DT 01-OCT-2003 (Tremblrel. 25, Last annotation update)  
 DE WW domain-containing oxidoreductase.  
 OS WWOX.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Kidney;  
 RA Strausberg R.;  
 RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=Mullerian duct includes surrounding region;  
 RX MEDLINE=22354683; PubMed=12466851;  
 RA The FANTOM Consortium,

RA the RIKEN Genome Exploration Research Group Phase I & II Team;  
 RT "Analysis of the mouse transcriptome based on functional annotation of  
 RL 60,770 full-length cDNAs."; J. Biol. Chem. 276:3361-3370(2001).  
 CC -|- SIMILARITY: BELONGS TO THE SHORT-CHAIN DEHYDROGENASES/REDUCTASES  
 CC (SDR) FAMILY.  
 DR EMBL; AF187014; AAF31693.1; -.  
 DR HSSP; Q13526; 1PIN.  
 DR MGD; MGI:1931237; Wwox.  
 DR GO; GO:0005739; C:mitochondrion; IDA.  
 DR GO; GO:0005634; C:nucleus; IDA.  
 DR GO; GO:0006917; P:induction of apoptosis; IDA.  
 DR InterPro; IPR002198; ADH short.  
 DR InterPro; IPR001202; WW\_Rsp5\_WWP.  
 DR Pfam; PF00106; adh short; 1.  
 DR Pfam; PF00397; WW\_2.  
 DR SMART; SM00456; WW; 2.  
 DR PROSITE; PS01159; WW DOMAIN 1; 2.  
 DR PROSITE; PS0020; WW DOMAIN 2; 2.  
 KW Oxidoreductase.  
 SQ SEQUENCE 414 AA; 46512 MW; 3C83AE3085B6A931 CRC64;

Query Match 94.0%; Score 2071; DB 11; Length 414;  
 Best Local Similarity 93.9%; Pred. No. 1.9e-165;  
 Matches 388; Conservative 10; Mismatches 15; Indels 0; Gaps 0;  
 QY 1 MAALRYAGLDGDDTSEDELPDGGWEERTTKDGVVYANHTTEKTQWEHPKTKRKRKRVAGDLP 60  
 DB 1 MAALRYAGLDGDDTSEDELPDGGWEERTTKDGVVYANHTTEKTQWEHPKTKRKRKRVAGDLP 60  
 QY 61 YGWEQETDENGQVFPVDHINKRTTYLDPRLAFTVDNDNPKPTTRQYDGSSTTAMEILQGR 120  
 DB 61 YGWEQETDENGQVFPVDHINKRTTYLDPRLAFTVDNDNPKPTTRQYDGSSTTAMEILQGR 120  
 QY 121 DFTGKVVVVTGANGSIGPETAKSPALHGAHVILACRNMARASEAVSRILEEWHKAKVEAM 180  
 DB 121 DFTGKVVVVTGANGSIGPETAKSPALHGAHVILACRNMARASEAVSRILEEWHKAKVEAM 180  
 QY 181 TLDLALLRSVOHEAFKAKVPLHVLVCNAATFALPWSLTQDGLTTFQVNHGHHFYLV 240  
 DB 181 TLDLAVLSVQHFABAFKAKVPLHVLVCNAGTFALPWGLTKDGLTTFQVNHGHHFYLV 240  
 QY 241 QLLQDVLCSAPARVIVVSSSHRFTDINDSLGKIDFSRLSPTKNDYWAMLAYNRSKLCN 300  
 DB 241 QLLQDVLCSAPARVIVVSSSHRFTDINDSGKLDLSRLSPRSDYWAMLAYNRSKLCN 300  
 QY 301 ILFSNELHRLSPRGVTSNAVHPGNMYSNTHRSWVYVTLFTLARPTKSMQQAATTV 360  
 DB 301 ILFSNELHRLSPRGVTSNAVHPGNMYSNTHRSWVYVTLFTLARPTKSMQQAATTV 360  
 QY 361 YCAAVPELEGGMVFNCCRCMPSPQAQSEETARTLWALSERLIOERLGSQS 413  
 DB 361 YCAVAPELEGGMVFNCCRCCLPSEEAQSEETARALWELSERLIOERLGSQS 413

RESULT 4  
 Q9JLF5 PRELIMINARY; PRT; 414 AA.  
 AC Q9JLF5;  
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)  
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE WW-domain oxidoreductase.  
 GN WWOX.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OC NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=21264809; PubMed=11058590;  
 RA Chang N.-S., Pratt N., Heath J., Schultz L., Sleeve D., Carey G.B.,  
 RA Zevotek N.;

RT "Hyaluronidase Induction of a WW Domain-Containing Oxidoreductase that  
 RT Enhances Tumor Necrosis Factor Cytotoxicity."; J. Biol. Chem. 276:3361-3370(2001).  
 CC -|- SIMILARITY: BELONGS TO THE SHORT-CHAIN DEHYDROGENASES/REDUCTASES  
 CC (SDR) FAMILY.  
 DR EMBL; AF187014; AAF31693.1; -.  
 DR HSSP; Q13526; 1PIN.  
 DR MGD; MGI:1931237; Wwox.  
 DR GO; GO:0005739; C:mitochondrion; IDA.  
 DR GO; GO:0005634; C:nucleus; IDA.  
 DR GO; GO:0006917; P:induction of apoptosis; IDA.  
 DR InterPro; IPR002198; ADH short.  
 DR InterPro; IPR001202; WW\_Rsp5\_WWP.  
 DR Pfam; PF00106; adh short; 1.  
 DR Pfam; PF00397; WW\_2.  
 DR SMART; SM00456; WW; 2.  
 DR PROSITE; PS01159; WW DOMAIN 1; 2.  
 DR PROSITE; PS0020; WW DOMAIN 2; 2.  
 KW Oxidoreductase.  
 SQ SEQUENCE 414 AA; 46484 MW; 1B8BA12882B6BE3P CRC64;  
 Query Match 93.8%; Score 2067; DB 11; Length 414;  
 Best Local Similarity 93.7%; Pred. No. 4.1e-165;  
 Matches 387; Conservative 10; Mismatches 16; Indels 0; Gaps 0;  
 QY 1 MAALRYAGLDGDDTSEDELPDGGWEERTTKDGVVYANHTTEKTQWEHPKTKRKRKRVAGDLP 60  
 DB 1 MAALRYAGLDGDDTSEDELPDGGWEERTTKDGVVYANHTTEKTQWEHPKTKRKRKRVAGDLP 60  
 QY 61 YGWEQETDENGQVFPVDHINKRTTYLDPRLAFTVDNDNPKPTTRQYDGSSTTAMEILQGR 120  
 DB 61 YGWEQETDENGQVFPVDHINKRTTYLDPRLAFTVDNDNPKPTTRQYDGSSTTAMEILQGR 120  
 QY 121 DFTGKVVVVTGANGSIGPETAKSPALHGAHVILACRNMARASEAVSRILEEWHKAKVEAM 180  
 DB 121 DFTGKVVVVTGANGSIGPETAKSPALHGAHVILACRNMARASEAVSRILEEWHKAKVEAM 180  
 QY 181 TLDLALLRSVOHEAFKAKVPLHVLVCNAATFALPWSLTQDGLTTFQVNHGHHFYLV 240  
 DB 181 TLDLAVLSVQHFABAFKAKVPLHVLVCNAGTFALPWGLTKDGLTTFQVNHGHHFYLV 240  
 QY 241 QLLQDVLCSAPARVIVVSSSHRFTDINDSLGKIDFSRLSPTKNDYWAMLAYNRSKLCN 300  
 DB 241 QLLQDVLCSAPARVIVVSSSHRFTDINDSGKLDLSRLSPRSDYWAMLAYNRSKLCN 300  
 QY 301 ILFSNELHRLSPRGVTSNAVHPGNMYSNTHRSWVYVTLFTLARPTKSMQQAATTV 360  
 DB 301 ILFSNELHRLSPRGVTSNAVHPGNMYSNTHRSWVYVTLFTLARPTKSMQQAATTV 360  
 QY 361 YCAAVPELEGGMVFNCCRCMPSPQAQSEETARTLWALSERLIOERLGSQS 413  
 DB 361 YCAVAPELEGGMVFNCCRCCLPSEEAQSEETARALWELSERLIOERLGSQS 413

RESULT 5  
 Q9NRF6 PRELIMINARY; PRT; 363 AA.  
 AC Q9NRF6;  
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)  
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
 DE For I protein.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OC NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=20320695; PubMed=10861292;  
 RA Ried K., Finnis M., Hobson L., Mangelsdorf M., Dayan S.,  
 RA Nancarrow J.K., Woolatt E., Kremmidiotis G., Gardner A., Venter D.,  
 RA Baker E., Richards R.I.;  
 RT "Common chromosomal fragile site FRA16D sequence: Identification of the

RT FOR gene spanning FRA16D and homozygous deletions and translocation  
 RT breakpoints in cancer cells.";  
 RL Hum. Mol. Genet. 9:1651-1663(2000).  
 CC -1- SIMILARITY: BELONGS TO THE SHORT-CHAIN DEHYDROGENASES/REDUCTASES  
 CC (SDR) FAMILY.

DR EMBL; AF227526; AAF82053.1; -.  
 DR HSSP; Q13526; 1PIN.  
 DR GO; GO:0016491; F:oxidoreductase activity; IEA.  
 DR GO; GO:0008152; P:metabolism; IEA.  
 DR InterPro; IPR002198; ADH short.  
 DR InterPro; IPR001202; WW\_ResP\_WWP.  
 DR Pfam; PF00106; adh\_short; 1.  
 DR Pfam; PF00397; WW; 2.  
 DR SMART; SM00456; WW; 2.  
 DR PROSITE; PS01159; WW\_DOMAIN\_1; 2.  
 DR PROSITE; PS00020; WW\_DOMAIN\_2; 2.  
 KW Oxidoreductase.  
 SQ SEQUENCE 363 AA; 41157 MW; C8DE29E7F11781D6 CRC64;

Query Match 84.8%; Score 1869; DB 4; Length 363;  
 Best Local Similarity 99.7%; Pred. No. 1.5e-148;  
 Matches 351; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MAALRYAGLDLDDSDDELPPGWEERTTKDGWVYANHTTEKTQWEHPKTKGKRVRAGDLP 60  
 |||||  
 DB 1 MAALRYAGLDLDDSDDELPPGWEERTTKDGWVYANHTTEKTQWEHPKTKGKRVRAGDLP 60  
 |||||  
 QY 61 YGWEQETDENGQVFFVDHINKRTTYLDPLAFTVDNDNPKPTTRQYDGTSTAMEILQGR 120  
 |||||  
 DB 61 YGWEQETDENGQVFFVDHINKRTTYLDPLAFTVDNDNPKPTTRQYDGTSTAMEILQGR 120  
 |||||  
 QY 121 DFTGKVVVVTGANGSIGFETAKSFALHGAHVILACRNMARASEAVSRILEEHWKAKVEAM 180  
 |||||  
 DB 121 DFTGKVVVVTGANGSIGFETAKSFALHGAHVILACRNMARASEAVSRILEEHWKAKVEAM 180  
 |||||  
 QY 181 TLDLALLRSVQHFAFAFAKAKNVLPLVLCNAATFALPWSLTKDGLTTTQVNHGHFYLV 240  
 |||||  
 DB 181 TLDLALLRSVQHFAFAFAKAKNVLPLVLCNAATFALPWSLTKDGLTTTQVNHGHFYLV 240  
 |||||  
 QY 241 QLLQDVLCRSAPARVIVVSSSHRFTDINDSLGKLDLDFSLPTKNDYMWAMLYNRSKLCN 300  
 |||||  
 DB 241 QLLQDVLCRSAPARVIVVSSSHRFTDINDSLGKLDLDFSLPTKNDYMWAMLYNRSKLCN 300  
 |||||  
 QY 301 ILFSNELHRLSPRGVTSNAVHPGNMYSNIHRSWMVYTLFTLARPPFTKSM 352  
 |||||  
 DB 301 ILFSNELHRLSPRGVTSNAVHPGNMYSNIHRSWMVYTLFTLARPPFTKSM 352  
 |||||

## RESULT 6

Q9D2B3 ID Q9D2B3 PRELIMINARY; PRT; 354 AA.  
 AC Q9D2B3;  
 DT 01-JUN-2001 (TREMELrel. 17, Created)  
 DT 01-JUN-2001 (TREMELrel. 17, Last sequence update)  
 DT 01-JUN-2003 (TREMELrel. 24, Last annotation update)  
 DE 5330426P09Rik protein.  
 GN WWOX OR 5330426P09RIK.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]

SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=Pituitary;  
 RX MEDLINE=21085660; PubMed=11217851;  
 RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,  
 RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,  
 RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,  
 RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,  
 RA Fleischmann W., Gaasterland T., Giasi C., King B., Kochiwa H.,  
 RA Kuehl P., Lewis S., Matsuo Y., Nikaide I., Pesole G., Quackenbush J.,  
 RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,

RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,  
 RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,  
 RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,  
 RA Gustincich S., Hall D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,  
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,  
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,  
 RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,  
 RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,  
 RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohsaki S.,  
 RA Hayashizaki Y.,  
 RT "Functional annotation of a full-length mouse cDNA collection.";  
 RL Nature 409:685-690(2001).  
 CC -1- SIMILARITY: BELONGS TO THE SHORT-CHAIN DEHYDROGENASES/REDUCTASES  
 CC (SDR) FAMILY.  
 DR EMBL; AK019911; BAB31911.1; -.  
 DR HSSP; Q13526; 1PIN.  
 DR MGD; MGI:1931237; Wwox.  
 DR GO; GO:0016491; F:oxidoreductase activity; IEA.  
 DR GO; GO:0008152; P:metabolism; IEA.  
 DR InterPro; IPR002198; ADH short.  
 DR InterPro; IPR001202; WW\_ResP\_WWP.  
 DR Pfam; PF00106; adh\_short; 1.  
 DR Pfam; PF00397; WW; 2.  
 DR SMART; SM00456; WW; 2.  
 DR PROSITE; PS01159; WW\_DOMAIN\_1; 2.  
 DR PROSITE; PS00020; WW\_DOMAIN\_2; 2.  
 KW Oxidoreductase.  
 SQ SEQUENCE 354 AA; 40018 MW; 41CBA77635E5E985 CRC64;

Query Match 81.0%; Score 1785; DB 11; Length 354;  
 Best Local Similarity 95.2%; Pred. No. 1.6e-141;  
 Matches 335; Conservative 8; Mismatches 9; Indels 0; Gaps 0;

QY 1 MAALRYAGLDLDDSDDELPPGWEERTTKDGWVYANHTTEKTQWEHPKTKGKRVRAGDLP 60  
 |||||  
 DB 1 MAALRYAGLDLDDSDDELPPGWEERTTKDGWVYANHTTEKTQWEHPKTKGKRVRAGDLP 60  
 |||||  
 QY 61 YGWEQETDENGQVFFVDHINKRTTYLDPLAFTVDNDNPKPTTRQYDGTSTAMEILQGR 120  
 |||||  
 DB 61 YGWEQETDENGQVFFVDHINKRTTYLDPLAFTVDNDNPKPTTRQYDGTSTAMEILQGR 120  
 |||||  
 QY 121 DFTGKVVVVTGANGSIGFETAKSFALHGAHVILACRNMARASEAVSRILEEHWKAKVEAM 180  
 |||||  
 DB 121 DFTGKVVVVTGANGSIGFETAKSFALHGAHVILACRNMARASEAVSRILEEHWKAKVEAM 180  
 |||||  
 QY 181 TLDLALLRSVQHFAFAFAKAKNVLPLVLCNAATFALPWSLTKDGLTTTQVNHGHFYLV 240  
 |||||  
 DB 181 TLDLALLRSVQHFAFAFAKAKNVLPLVLCNAATFALPWSLTKDGLTTTQVNHGHFYLV 240  
 |||||  
 QY 241 QLLQDVLCRSAPARVIVVSSSHRFTDINDSLGKLDLDFSLPTKNDYMWAMLYNRSKLCN 300  
 |||||  
 DB 241 QLLQDVLCRSAPARVIVVSSSHRFTDINDSLGKLDLDFSLPTKNDYMWAMLYNRSKLCN 300  
 |||||  
 QY 301 ILFSNELHRLSPRGVTSNAVHPGNMYSNIHRSWMVYTLFTLARPPFTKSM 352  
 |||||  
 DB 301 ILFSNELHRLSPRGVTSNAVHPGNMYSNIHRSWMVYTLFTLARPPFTKSM 352  
 |||||

## RESULT 7

Q9D339 ID Q9D339 PRELIMINARY; PRT; 367 AA.  
 AC Q9D339;  
 DT 01-JUN-2001 (TREMELrel. 17, Created)  
 DT 01-JUN-2001 (TREMELrel. 17, Last sequence update)  
 DT 01-JUN-2003 (TREMELrel. 24, Last annotation update)  
 DE 9030416C10Rik protein.  
 GN WWOX OR 9030416C10RIK.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.

RC STRAIN=CS7BL/6J; TISSUE=Colon;  
 RX MEDLINE=21085660; PubMed=11217851;  
 RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,  
 RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,  
 RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,  
 RA Saito T., Okazaki Y., Gojohori T., Bono H., Kasukawa T., Saito R.,  
 RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,  
 RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,  
 RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,  
 RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,  
 RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,  
 RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,  
 RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,  
 RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kaniya M., Lee N.H.,  
 RA Lyons P., Marchionni L., Mashima J., Marzarelli J., Mombaerts P.,  
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,  
 RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-P.,  
 RA Suzuki H., Iyoy-oka K., Wang K.H., Weitz C., Whitaker C., Wilming L.,  
 RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohetsuki S.,  
 RA Hayashizaki Y.;  
 RT "Functional annotation of a full-length mouse cDNA collection.";  
 RL Nature 409:685-690(2001).  
 CC -!- SIMILARITY: BELONGS TO THE SHORT-CHAIN DEHYDROGENASES/REDUCTASES  
 CC (SDR) FAMILY.  
 DR EMBL; AK018507; BAB31244.1; -.  
 DR HSSP; Q13526; 1PIN.  
 DR MGD; MGI:1931237; Wwox.  
 DR GO; GO:0016491; F:oxidoreductase activity; IEA.  
 DR GO; GO:0008152; P:metabolism; IEA.  
 DR InterPro; IPR002198; ADH\_short.  
 DR InterPro; IPR001202; WW Rep5\_WWP.  
 DR Pfam; PF00106; adh\_short; 1.  
 DR Pfam; PF00397; WW; 2.  
 DR SMART; SM00456; WW; 2.  
 DR PROSITE; PS01159; WW DOMAIN 1; 2.  
 DR PROSITE; PS00020; WW DOMAIN\_2; 2.  
 DR Oxidoreductase.  
 KW Oxidoreductase.  
 SQ SEQUENCE 367 AA; 41499 MW; B29A368E793B2C0F CRC64;

Query Match 80.7%; Score 1777; DB 11; Length 367;  
 Best Local Similarity 94.6%; Pred. No. 7.9e-141; Indels 0; Gaps 0;  
 Matches 334; Conservative 9; Mismatches 10;  
 Qy 1 MAALRYAGLDDTDSDELPPGWEERTTKDGVVYVYANHEEKTQWEHPTGKRKRVAGDLP 60  
 Db 1 MAALRYAGLDDTDSDELPPGWEERTTKDGVVYVYANHEEKTQWEHPTGKRKRVAGDLP 60  
 Qy 61 YGWEQETDENGQVFVDHINKRTTYLDPRLAFTVDDNPTKPTTQRYDGSSTAMEILQGR 120  
 Db 61 YGWEQETDENGQVFVDHINKRTTYLDPRLAFTVDDNPTKPTTQRYDGSSTAMEILQGR 120  
 Qy 121 DFTGKVVVVVTGANGSIGFETAKSPALHGAHVILACRNARASEAVSRILEEWHKAKVEAM 180  
 Db 121 DFTGKVVVVVTGANGSIGFETAKSPALHGAHVILACRNARASEAVSRILEEWHKAKVEAM 180  
 Qy 181 TLDLALRSVQHFAEAFKAKNVPLHVLVCNAATFALPWSLTGKLETTFQVNHILGHFYL 240  
 Db 181 TLDLALRSVQHFAEAFKAKNVPLHVLVCNAATFALPWSLTGKLETTFQVNHILGHFYL 240  
 Qy 241 QLLODVLCRSAPARVIVVSSSHRFTDINDSLGKLDPSRLSPTKNDYVWMLAYNRSLCN 300  
 Db 241 QLLODVLCRSAPARVIVVSSSHRFTDINDSLGKLDPSRLSPTKNDYVWMLAYNRSLCN 300  
 Qy 301 ILFNSNELHRLSPRGVTSNAVHPGNMYSNTHRSWVYVYLLFTLARPFTKSMQ 353  
 Db 301 ILFNSNELHRLSPRGVTSNAVHPGNMYSNTHRSWVYVYLLFTLARPFTKSMK 353

RESULT 8

Q803A8

ID Q803A8

AC Q803A8

DT 01-JUN-2003 (TReMBLrel. 24, Created)

DT 01-JUN-2003 (TReMBLrel. 24, Last sequence update)  
 DT 01-OCT-2003 (TReMBLrel. 25, Last annotation update)  
 DE Similar to WW domain containing oxidoreductase.  
 OS Brachydanio rerio (Zebrafish) (Danio rerio).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Actinopterygii; Neopterygii; Teleostei; Osteichthyes; Cypriniformes;  
 OC Cyprinidae; Danio.  
 OX NCBI\_TaxID=7955;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Strausberg R.;  
 RL Submitted (JAN-2003) to the EMBL/GenBank/DDBJ databases.  
 DR EMBL; BC044560; AAH44560.1; -.  
 DR GO; GO:0016491; F:oxidoreductase activity; IEA.  
 DR GO; GO:0008152; P:metabolism; IEA.  
 DR InterPro; IPR002198; ADH\_short.  
 DR InterPro; IPR001202; WW Rep5\_WWP.  
 DR Pfam; PF00106; adh\_short; 1.  
 DR Pfam; PF00397; WW; 2.  
 DR SMART; SM00456; WW; 2.  
 DR PROSITE; PS01159; WW DOMAIN 1; 1.  
 DR PROSITE; PS00020; WW DOMAIN\_2; 2.  
 DR PROSITE; PS00020; WW DOMAIN\_2; 2.  
 SQ SEQUENCE 412 AA; 46322 MW; D4D10CA809138B8D CRC64;  
 Query Match 74.3%; Score 1637.5; DB 13; Length 412;  
 Best Local Similarity 72.8%; Pred. No. 4.9e-129;  
 Matches 297; Conservative 51; Mismatches 57; Indels 3; Gaps 1;  
 Qy 1 MAALRYAGLDDTDSDELPPGWEERTTKDGVVYVYANHEEKTQWEHPTGKRKRVAGDLP 60  
 Db 1 MAALRYAGLDDTDSDELPPGWEERTTKDGVVYVYANHEEKTQWEHPTGKRKRVAGDLP 60  
 Qy 61 YGWEQETDENGQVFVDHINKRTTYLDPRLAFTVDDNPTKPTTQRYDGSSTAMEILQGR 120  
 Db 61 YGWEQETDENGQVFVDHINKRTTYLDPRLAFTVDDNPTKPTTQRYDGSSTAMEILQGR 120  
 Qy 121 DFTGKVVVVVTGANGSIGFETAKSPALHGAHVILACRNARASEAVSRILEEWHKAKVEAM 180  
 Db 121 DFTGKVVVVVTGANGSIGFETAKSPALHGAHVILACRNARASEAVSRILEEWHKAKVEAM 180  
 Qy 181 TLDLALRSVQHFAEAFKAKNVPLHVLVCNAATFALPWSLTGKLETTFQVNHILGHFYL 240  
 Db 181 TLDLALRSVQHFAEAFKAKNVPLHVLVCNAATFALPWSLTGKLETTFQVNHILGHFYL 240  
 Qy 241 QLLODVLCRSAPARVIVVSSSHRFTDINDSLGKLDPSRLSPTKNDYVWMLAYNRSLCN 300  
 Db 241 QLLODVLCRSAPARVIVVSSSHRFTDINDSLGKLDPSRLSPTKNDYVWMLAYNRSLCN 300  
 Qy 301 ILFNSNELHRLSPRGVTSNAVHPGNMYSNTHRSWVYVYLLFTLARPFTKSMQ 360  
 Db 301 ILFNSNELHRLSPRGVTSNAVHPGNMYSNTHRSWVYVYLLFTLARPFTKSMQ 360  
 Qy 361 YCAVPELEGIGGMVFNCCRCMPSPQAQSEETARTLWALSERLIQGR 408  
 Db 361 YCAVPELEGIGGMVFNCCRCMPSPQAQSEETARTLWALSERLIQGR 408  
 Qy 358 YCAVPELEGIGGMVFNCCRCMPSPQAQSEETARTLWALSERLIQGR 405  
 Db 358 YCAVPELEGIGGMVFNCCRCMPSPQAQSEETARTLWALSERLIQGR 405

RESULT 9

Q9BTT8

ID Q9BTT8

AC Q9BTT8

DT 01-JUN-2001 (TReMBLrel. 17, Created)

DT 01-JUN-2001 (TReMBLrel. 17, Last sequence update)

DT 01-JUN-2003 (TReMBLrel. 24, Last annotation update)

DE Similar to WW domain-containing oxidoreductase (WwoxDelta6-8).

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OX NCBI\_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=Lung;

RA Strausberg R.;

Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.  
[2]  
RN SEQUENCE FROM N.A.  
RA Bednarek A.K., Keck-Waggoner C.L., Daniel R.L., Lafin K.J.,  
RA Kiguchi K., Brenner A.J., Aldaz C.M.;  
RT "WBOX, the FRA16D gene, behaves as a suppressor of tumor growth.";  
RN Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.  
[3]  
RN SEQUENCE FROM N.A.  
RA Kalline N., Chen X., Rolfs A., Halleck A., Hines L., Eisenstein S.,  
RA Koundinya M., Raphael J., Moreira D., Kelley T., LaBaer J., Lin Y.,  
RA Phelan M., Farmer A.;  
RT "Cloning of human full-length CDSS in BD Creator(TM) System Donor  
vector.";  
RN Submitted (MAY-2003) to the EMBL/GenBank/DBJ databases.  
CC -!- SIMILARITY: BELONGS TO THE SHORT-CHAIN DEHYDROGENASES/REDUCTASES  
(SDR) FAMILY  
DR EMBL; BC003184; AA03184.1; -;  
DR EMBL; AF395123; AAK81727.1; -;  
DR EMBL; BT007445; AAP36113.1; -;  
DR HSSP; Q13526; 1PIN.  
DR GO; GO:0016491; P:oxidoreductase activity; IEA.  
DR GO; GO:0008152; P:metabolism; IEA.  
DR InterPro; IPR002198; ADH\_short.  
DR Pfam; PF00106; adh\_short; WW.  
DR Pfam; PF00397; WW; 2.  
DR SMART; SM00456; WW; 2.  
DR PROSITE; PS01159; WW\_DOMAIN\_1; 2.  
DR PROSITE; PS00020; WW\_DOMAIN\_2; 2.  
KW Oxidoreductase.  
SQ SEQUENCE 234 AA; 26148 MW; C69FB63B87635F4 CRC64;  
  
Query Match 52.5%; Score 1157; DB 4; Length 234;  
Best Local Similarity 56.5%; Pred. No. 5e-89;  
Matches 234; Conservative 0; Mismatches 0; Indels 180; Gaps 1;  
  
QY 1 MAALRYAGLDLTDSELPPEWEERTTKDGWVYVYANHTTEKTQWEHPKTKGKRVAGDLP 60  
Db 1 MAALRYAGLDLTDSELPPEWEERTTKDGWVYVYANHTTEKTQWEHPKTKGKRVAGDLP 60  
  
QY 61 YWQEQTDENGQVFFVDHINKETTYLDPLAFTVDNDPKPTTRDYDGSITAMEILQGR 120  
Db 61 YWQEQTDENGQVFFVDHINKETTYLDPLAFTVDNDPKPTTRDYDGSITAMEILQGR 120  
  
QY 121 DFTGKVVVVTGANSIGFTAKSEALHGAHVILACRNMARASEAVSRIILEEWHKAKVEAM 180  
Db 121 DFTGKVVVVTGANSIGFTAKSEALHGAHVILACRNMARASEAVSRIILEEWHKAKVEAM 172  
  
QY 181 TLDLALLRSVQHFABFAKAKNPVLHVCNAATFALPWSLTQDGLTTPQVNLGHFYLIV 240  
Db 173 ----- 172  
  
QY 241 QLLQDVLCESAPARVIVVSSSHRFTDINDSLGKDFSLPTKNDYMWALNRSKLCN 300  
Db 173 ----- 172  
  
QY 301 ILFSLNELHRLSPRGVTSNAVHPGNMYSNIHRSWWVYTLFTLARPFPTKSMOOGAATTV 360  
Db 173 ----- 172  
  
QY 361 YCAAVELEGLOGMYFNCCRCMPSEPAQSEETARTLWALSERLIQERLGSQSG 414  
Db 181 YCAAVELEGLOGMYFNCCRCMPSEPAQSEETARTLWALSERLIQERLGSQSG 234  
  
RESULT 10  
Q920Y2  
ID Q920Y2 PRELIMINARY; PRT; 242 AA.  
AC Q920Y2;  
DT 01-DEC-2001 (T-EMBLrel. 19, Created)  
DT 01-DEC-2001 (T-EMBLrel. 19, Last sequence update)  
DT 01-OCT-2003 (T-EMBLrel. 25, Last annotation update)

DE WW-domain oxidoreductase (Fragment).  
GN WBOX.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Krummel K.A., Denison S.R., Calhoun E., Smith D.I.;  
RT "The Common Fragile Site FRA16D and its Associated Gene WBOX are  
RT Highly Conserved in Mouse at Fra8E1.";  
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AY046556; AAL03972.1; -;  
DR EMBL; AY046553; AAL03972.1; JOINED.  
DR EMBL; AY046554; AAL03972.1; JOINED.  
DR EMBL; AY046555; AAL03972.1; JOINED.  
DR MGD; MGI:1931237; Wbox.  
DR GO; GO:0005739; C:mitochondrion; IEA.  
DR GO; GO:0005634; C:nucleus; IEA.  
DR GO; GO:0006917; P:induction of apoptosis; IEA.  
FT NON TER 1  
SQ SEQUENCE 242 AA; 27036 MW; 35F91C09F8D815DA CRC64;  
  
Query Match 52.2%; Score 1149; DB 11; Length 242;  
Best Local Similarity 90.5%; Pred. No. 2.5e-88;  
Matches 218; Conservative 7; Mismatches 16; Indels 0; Gaps 0;  
  
QY 173 HKAKVEAMTLDLALLRSVQHFABFAKAKNPVLHVCNAATFALPWSLTQDGLTTPQV 232  
Db 1 HKAKVEAMTLDLALLRSVQHFABFAKAKNPVLHVCNAATFALPWSLTQDGLTTPQV 60  
  
QY 233 HLGHFYLVQLLQDVLCRSAPARVIVVSSSHRFTDINDSLGKDFSLPTKNDYMWALIA 292  
Db 61 HLGHFYLVQLLQDVLCRSAPARVIVVSSSHRFTDINDSLGKDFSLPTKNDYMWALIA 120  
  
QY 293 YNRSKLCNLFSLNELHRLSPRGVTSNAVHPGNMYSNIHRSWWVYTLFTLARPFPTKSM 352  
Db 121 YNRSKLCNLFSLNELHRLSPRGVTSNAVHPGNMYSNIHRSWWVYTLFTLARPFPTKSM 180  
  
QY 353 OOGAATTVYCAVPELEGLOGMYFNCCRCMPSEPAQSEETARTLWALSERLIQERLGSQ 412  
Db 181 OOGAATTVYCAVPELEGLOGMYFNCCRCMPSEPAQSEETARTLWALSERLIQERLGSQ 240  
  
QY 413 S 413  
Db 241 S 241  
  
RESULT 11  
Q9VLUS  
ID Q9VLUS PRELIMINARY; PRT; 409 AA.  
AC Q9VLUS;  
DT 01-MAY-2000 (T-EMBLrel. 13, Created)  
DT 01-MAY-2000 (T-EMBLrel. 13, Last sequence update)  
DT 01-JUN-2003 (T-EMBLrel. 24, Last annotation update)  
DE CG7221 protein (LD03827p).  
GN CG7221.  
OS Drosophila melanogaster (Fruit fly).  
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
OC Ephydroidea; Drosophilidae; Drosophila.  
OX NCBI TaxID=7227;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Berkely;  
RX MEDLINE=20196006; PubMed=10731132;  
RA Adams M.D., Celnikier S.B., Holt R.A., Evans C.A., Gocayne J.D.,  
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,  
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,  
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,  
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,  
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,

RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,  
RA Borokova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,  
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,  
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,  
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,  
RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,  
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,  
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,  
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,  
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,  
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,  
RA Palazolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,  
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E., Wang X.,  
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissenbach J.,  
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
RT "The genome sequence of *Drosophila melanogaster*.";  
RL Science 287:2185-2195 (2000).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Berkeley;  
RA Stapleton M., Brokstein P., Hong L., Aghayani A., Carlson J.,  
RA Champe W., Chavez C., Dorsett V., Dresnek D., Farfan D., Frise E.,  
RA George R., Gonzalez M., Guarin H., Grommler B., Li P., Liao G.,  
RA Miranda A., Mungall C.J., Nunoo J., Pacleb J., Paragas V., Park S.,  
RA Patel S., Phouanavong S., Wan K., Yu C., Lewis S.E., Rubin G.M.,  
RA Celniker S.;  
RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.  
CC -!- SIMILARITY: BELONGS TO THE SHORT-CHAIN DEHYDROGENASES/REDUCTASES  
(SDR) FAMILY.  
DR EMBL; AE003619; AAF52587.1; -;  
DR EMBL; AY119574; AAM50228.1; -;  
DR HSSP; Q13526; 1PIN.  
DR FlyBase; FBgn0031972; CG7221.  
DR GO; GO:0016491; F:oxidoreductase activity; IEA.  
DR GO; GO:0008152; P:metabolism; IEA.  
DR InterPro; IPR002198; ADH short.  
DR InterPro; IPR001202; WW\_Rsp5\_WWP.  
DR Pfam; PF00106; adh\_short; 1.  
DR Pfam; PF00397; WW\_2.  
DR PROSITE; PS50020; WW\_DOMAIN\_2; 2.  
DR Oxidoreductase.  
SQ SEQUENCE 409 AA; 46510 MW; 7893BF4C39A7454C CRC64;

Query Match 46.4%; Score 1023; DB 5; Length 409;  
Best Local Similarity 49.5%; Pred. No. 2e-77;  
Matches 200; Conservative 68; Mismatches 126; Indels 10; Gaps 6;

QY 9 LDDTSDSELPFGMEERTTKDGVVYANHTBEKTOWEHPKTKGRKRVAGDLPYGWEQETD 68  
Db 4 LPDTSDELPPGWEERATDDGTVCVYVQOQKTSQWHTPTGRSKRIITGELPLGWEKYD 63

QY 69 ENGQVF-FVDHINKRTYLDRLAFTVDNDPTK--PTTRQRYDGSSTAMEILQGRDFTGK 125  
Db 64 EQGRFMPINKETQRTNVDPRLAFVVEE-PTQNVAVQRQFDSCTALQVLHGKDLHGR 122

QY 126 VVVVTGANGSGIGETAKSFAHGAHVILACRNAPASAVSRILLEEHWKAVEA--MTLD 183  
Db 123 TALITGANGCIGYETARSLAHHGCEIIFACRRNSAEAAIRIAQERPAARSRCKFAALD 182

QY 184 LALLRSVQHFAEAFKAKNVPLHLVLCNAATFALPWSLTNDGLTTFQVNHHLGHFYLVL 243  
Db 183 LSSLSRVQRFVEEIKQSVSHIDYLLNLNAGVPALPYTRTVTVDGLTTFQVSHLSHFYLTQL 242

QY 244 QDVLCRSAPARVIVSSSHRFTDINDSLGKLDLFRSLSPTRKNDYWLAMLAYNRSKLNILF 303  
Db 243 ETLF--DYKTRIIIVLSSSHRFAUL--EVENLAVHLSPPPEKYWSMMAYNNAKLNVLF 298

QY 304 SNEHLRRLSPRGVTSNAVHPGNMYSNTHRSWWVYTLFTLARPPTKSMQCGAATTVYCA 363  
Db 299 AQELAQRWKQGISVFSLHPGNMSSDLRSNFWFVRLFLFAIVRPPTKSLQQAASIIYCA 358

QY 364 APELEGGLGMYFNCCRCMPSEPAQSEETARTLWALSERLQIE 407  
Db 359 TANELTGLSLGYFNCCFCEPSKLSKSAALQQWLKLSENLIAE 402

RESULT 12  
Q9NRK1 PRELIMINARY; PRT; 191 AA.  
AC Q9NRK1;  
DT 01-OCT-2000 (TrEMBLrel. 15, Created)  
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
DE Fragile 16D oxido reductase (Fragment).  
GN FOR.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=20320695; PubMed=10861292;  
RA Ried K., Finnis M., Hobson L., Mangelsdorf M., Dayan S.,  
RA Nancarrow J.K., Woolatt E., Kremmidiotis G., Gardner A., Venter D.,  
RA Baker E., Richards R.I.;  
RT "Common chromosomal fragile site FRA16D sequence: Identification of the  
RT FOR gene spanning FRA16D and homozygous deletions and translocation  
RT breakpoints in cancer cells.";  
RL Hum. Mol. Genet. 9:1651-1663 (2000).  
DR EMBL; AF217430; AAF78197.1; -;  
DR EMBL; AF217491; AAF78197.1; JOINED.  
DR EMBL; AF217492; AAF78197.1; JOINED.  
DR GO; GO:0016491; F:oxidoreductase activity; NAS.  
FT NON TER 1  
SQ SEQUENCE 191 AA; 21721 MW; 030FC7E106D86B55 CRC64;

Query Match 42.8%; Score 942; DB 4; Length 191;  
Best Local Similarity 99.4%; Pred. No. 4.3e-71;  
Matches 179; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 173 HKAKVEAMTDLALLRSVQHFAEAFKAKNVPLHLVLCNAATFALPWSLTNDGLTTFQV 232  
Db 1 HKAKVETMTDLALLRSVQHFAEAFKAKNVPLHLVLCNAATFALPWSLTNDGLTTFQV 60

QY 233 HLGHFYLVLQDLVLCRSAPARVIVSSSHRFTDINDSLGKLDLFRSLSPTRKNDYWLAMLA 292  
Db 61 HLGHFYLVLQDLVLCRSAPARVIVSSSHRFTDINDSLGKLDLFRSLSPTRKNDYWLAMLA 120

QY 293 YNRSKLCNLFESNEHLRRLSPRGVTSNAVHPGNMYSNTHRSWWVYTLFTLARPPTKSM 352  
Db 121 YNRSKLCNLFESNEHLRRLSPRGVTSNAVHPGNMYSNTHRSWWVYTLFTLARPPTKSM 180

RESULT 13  
Q9NPC9 PRELIMINARY; PRT; 189 AA.  
AC Q9NPC9;  
DT 01-OCT-2000 (TrEMBLrel. 15, Created)  
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
DE FOR III protein (Truncated WW-domain oxidoreductase) (WW domain-  
DE containing oxidoreductase isoform FORIII).

```
GN WWOX.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20320695; PubMed=10861292;
RA Ried K., Finniss M., Hobson L., Mangelsdorf M., Dayan S., Venter D.,
RA Nancarrow J.K., Woolatt E., Kremmidiotis G., Gardner A.,
RA Baker E., Richards R.I.;
RT "Common chromosomal fragile site FRA16D sequence: Identification of the
RT FOR gene spanning FRA16D and homozygous deletions and translocation
RT breakpoints in cancer cells.";
RL Hum. Mol. Genet. 9:1651-1663(2000).
RN [2]
RP SEQUENCE FROM N.A.
RX PubMed=11572989;
RA Chang N.-S., Pratt N., Carey G., Heath J.;
RA "A Novel WW Domain Oxidoreductase Enhances Tumor Necrosis Factor-
RT Mediated Cell Death and is a Partner of p53 in Apoptosis.";
RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RX PubMed=11572989;
RA Paige A.J.W., Taylor K.J., Taylor C., Hillier S.G., Farrington S.,
RA Scott D., Porteous D.J., Smyth J.F., Gabra H., Watson J.E.V.;
RT "WWOX: A candidate tumor suppressor gene involved in multiple tumor
RT types.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:11417-11422(2001).
CC -1- SIMILARITY: BELONGS TO THE SHORT-CHAIN DEHYDROGENASES/REDUCTASES
CC (SDR) FAMILY.
DR EMBL; AF227528; AAF82055.1; -.
DR EMBL; AF187015; AAF31694.1; -.
DR EMBL; AF325429; AAL05450.1; -.
DR EMBL; AF325423; AAL05450.1; JOINED.
DR EMBL; AF325424; AAL05450.1; JOINED.
DR EMBL; AF325425; AAL05450.1; JOINED.
DR EMBL; AF325426; AAL05450.1; JOINED.
DR EMBL; AF325427; AAL05450.1; JOINED.
DR HSP; Q13526; 1PIN.
DR GO; GO:0016491; F:oxidoreductase activity; NAS.
DR InterPro; IPR002198; ADH_Short.
DR InterPro; IPR001202; WW_Rep5_WWP.
DR Pfam; PF00106; adh_short; 1.
DR Pfam; PF00397; WW; 2.
DR SMART; SM00456; WW; 2.
DR PROSITE; PS01159; WW_DOMAIN_1; 2.
DR PROSITE; PSS0020; WW_DOMAIN_2; 2.
KW Oxidoreductase.
SQ SEQUENCE 189 AA; 21559 MW; 82155A9AD7C924C7 CRC64;

Query Match 42.1%; Score 927; DB 4; Length 189;
Best Local Similarity 100.0%; Pred. No. 7.7e-70;
Matches 172; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MAALRYAGLDDTDSDELPGWEERTTKDGVVYANHTTEKTQWEHPKTKRKRKRGADLP 60
Db 1 MAALRYAGLDDTDSDELPGWEERTTKDGVVYANHTTEKTQWEHPKTKRKRKRGADLP 60

Qy 61 YGWEQETDENGQVFFVDHINKRTTYLDPRLAFTVDDNPTKPTTRQRYDGSSTTAMEILQGR 120
Db 61 YGWEQETDENGQVFFVDHINKRTTYLDPRLAFTVDDNPTKPTTRQRYDGSSTTAMEILQGR 120

Qy 121 DFTGKVVVVTGANGSGIGFETAKSFALHGAHVILACRNARASEAVSRILEW 172
Db 121 DFTGKVVVVTGANGSGIGFETAKSFALHGAHVILACRNARASEAVSRILEW 172

RESULT 14
Q96KM3 ID Q96KM3 PRELIMINARY; PRT; 213 AA.
AC Q96KM3
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
```

```
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Aberrant WW domain-containing oxidoreductase.
GN WWOX.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX PubMed=11572989;
RA Paige A.J.W., Taylor K.J., Taylor C., Hillier S.G., Farrington S.,
RA Scott D., Porteous D.J., Smyth J.F., Gabra H., Watson J.E.V.;
RT "WWOX: A candidate tumor suppressor gene involved in multiple tumor
RT types.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:11417-11422(2001).
DR EMBL; AF325432; AAL05451.1; -.
DR EMBL; AF325423; AAL05451.1; JOINED.
DR EMBL; AF325424; AAL05451.1; JOINED.
DR EMBL; AF325425; AAL05451.1; JOINED.
DR EMBL; AF325426; AAL05451.1; JOINED.
DR EMBL; AF325433; AAL05451.1; JOINED.
DR InterPro; IPR001202; WW_Rep5_WWP.
DR Pfam; PF00397; WW; 2.
DR SMART; SM00456; WW; 2.
DR PROSITE; PS01159; WW_DOMAIN_1; 2.
DR PROSITE; PSS0020; WW_DOMAIN_2; 2.
SQ SEQUENCE 213 AA; 23868 MW; A21054FF8214CC7C CRC64;

Query Match 34.1%; Score 750.5; DB 4; Length 213;
Best Local Similarity 79.7%; Pred. No. 6.2e-55;
Matches 145; Conservative 5; Mismatches 7; Indels 25; Gaps 2;

Qy 1 MAALRYAGLDDTDSDELPGWEERTTKDGVVYANHTTEKTQWEHPKTKRKRKRGADLP 60
Db 1 MAALRYAGLDDTDSDELPGWEERTTKDGVVYANHTTEKTQWEHPKTKRKRKRGADLP 60

Qy 61 YGWEQETDENGQVFFVDHINKRTTYLDPRLAFTVDDNPTKPTTRQRYDGSSTTAMEILQGR 120
Db 61 YGWEQETDENGQVFFVDHINKRTTYLDPRLAFTVDDNPTKPTTRQRYDGSSTTAMEILQGR 120

Qy 121 DFTGKVVVVTGANGSGIGFETAKSFALHGAHVILACRNARASEAVSRILEW 180
Db 121 DFTGKVVVVTGANGSGIGFETAKSFALHGAHVILACRNARASEAVSRILEW 180

Qy 181 TL 182
Db 156 SL 157

RESULT 15
Q96RF2 ID Q96RF2 PRELIMINARY; PRT; 311 AA.
AC Q96RF2
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE WWOXdelta5-8
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX Bednarek A.K., Keck-Waggoner C.L., Daniel R.L., Laflin K.J.,
RA Kiguchi K., Brenner A.J., Aldaz C.M.;
RT "WWOX, the FRA16D gene, behaves as a suppressor of tumor growth.";
RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF395124; AAK81728.1; -.
DR InterPro; IPR001202; WW_Rep5_WWP.
DR Pfam; PF00397; WW; 2.
DR SMART; SM00456; WW; 2.
DR PROSITE; PS01159; WW_DOMAIN_1; 2.
```





GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model  
Run on: May 24, 2004, 21:49:29 ; Search time 8833 Seconds  
(without alignments)  
11109.321 Million cell updates/sec  
Title: US-09-978-318B-2  
Perfect score: 2264  
Sequence: 1 gcagcgccagcgagcg.....aaaaaaaaaaaaaaaa 2264

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0  
Searched: 3470272 seqs, 21671516995 residues  
Total number of hits satisfying chosen parameters: 5940544

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :		GenEmbl:*	
1:	gb_ba:*	21	567.2
2:	gb_htg:*	22	518.2
3:	gb_in:*	23	516.6
4:	gb_om:*	C 24	483.6
5:	gb_ov:*	C 25	483.6
6:	gb_pat:*	C 26	433.6
7:	gb_ph:*	C 27	416.8
8:	gb_pl:*	28	342.4
9:	gb_pr:*	29	280.6
10:	gb_ro:*	30	270.4
11:	gb_sts:*	31	266
12:	gb_sy:*	32	266
13:	gb_un:*	33	266
14:	gb_vi:*	34	249.4
15:	em_ba:*	C 35	237.8
16:	em_fun:*	36	235.2
17:	em_hum:*	37	232
18:	em_in:*	38	232
19:	em_mu:*	C 39	232
20:	em_om:*	C 40	232
21:	em_or:*	C 41	211.6
22:	em_ov:*	C 42	211.6
23:	em_pat:*	C 43	211.6
		44	211
		45	206.8

ALIGNMENTS

RESULT 1  
AF211943  
LOCUS AF211943 2264 bp mRNA linear PRI 05-MAY-2000  
DEFINITION Homo sapiens WW domain-containing protein WWOX mRNA, complete cds.  
ACCESSION AF211943  
VERSION AF211943.1 GI:6729682  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 2264)  
AUTHORS Bednarek,A.K., Lafin,K.J., Daniel,R.L., Liao,Q., Hawkins,K.A. and Aldaz,C.M.  
TITLE WWOX, a novel WW domain-containing protein mapping to human

Pred. No. is the number of results predicted by chance to have a







1 CAGTGCAGGCGGTGAGCGGTGCGGCCCGCAGCGCGCGGGTCTCTGTTTGGAGCGGAGT 60  
62 GAGTTCTGAGCGAGTGGACCCCGGAGCGGGCGATAGGGGGGCGAGGTGCTTCCACAGTC 121  
61 GAGTTCTGAGCGAGTGGACCCCGGAGCGGGCGATAGGGGGGCGAGGTGCTTCCACAGTC 120  
122 AGCCATGCGAGCGCTCGCTTACCGCGGGCTTGGACGACACGAGCAGTGGAGCAGAGTGGCC 181  
121 AGCCATGCGAGCGCTCGCTTACCGCGGGCTTGGACGACACGAGCAGTGGAGCAGAGTGGCC 180  
182 TCCGGGCTGGAGGAGAGAAACCAAGGACGCGTGGGTTTACTAGCCCAATCACACCGA 241  
181 TCCGGGCTGGAGGAGAGAAACCAAGGACGCGTGGGTTTACTAGCCCAATCACACCGA 240  
242 GGAAGAAGCTCAGTGGGAACATCCAAAACCTGAAAAGGAAAACGAGTGGCAGAGATTT 301  
241 GGAAGAAGCTCAGTGGGAACATCCAAAACCTGAAAAGGAAAACGAGTGGCAGAGATTT 300  
302 GCCATACGGATGGAAACAGAAACTGATGAGAACGGACAGTGTGTTTGTTCACCATAT 361  
301 GCCATACGGATGGAAACAGAAACTGATGAGAACGGACAGTGTGTTTGTTCACCATAT 360  
362 AATAAAGAACCACTTCTGGACCCAGACTGGCGTTTACTGTGATGATATATCCGAC 421  
361 AATAAAGAACCACTTCTGGACCCAGACTGGCGTTTACTGTGATGATATATCCGAC 420  
422 CAAGCCAAACCCGCGAAGATACGACGCGACGACCACTGGCATGGAAATTTCTCAGGG 481  
421 CAAGCCAAACCCGCGAAGATACGACGCGACGACCACTGGCATGGAAATTTCTCAGGG 480  
482 CCGGGATTTCACTGGCAAGTGGTGTGGTCACTGGAGCTAAATCAGGAATAGGGTTGGA 541  
481 CCGGGATTTCACTGGCAAGTGGTGTGGTCACTGGAGCTAAATCAGGAATAGGGTTGGA 540  
542 AACCGCAAGTCTTTTGGCCCTCCATGGTGCACATGTGATCTTGGCCTGCAGGAACATGGC 601  
541 AACCGCAAGTCTTTTGGCCCTCCATGGTGCACATGTGATCTTGGCCTGCAGGAACATGGC 600  
602 AAGGGGAGTGAAGCAGTGTACGCAATTTAGAGAAATGSCATAAAGCCAGGTAGAGC 661  
601 AAGGGGAGTGAAGCAGTGTACGCAATTTAGAGAAATGSCATAAAGCCAGGTAGAGC 660  
662 AATGACCTGGACCTCGCTCGTCCGTAGCGGTGACAGATTTTCTGTAAGCATTTCAAGGC 721  
661 AATGACCTGGACCTCGCTCGTCCGTAGCGGTGACAGATTTTCTGTAAGCATTTCAAGGC 720  
722 CAAGAATGTGCTCTTCAATGTGCTTGTGTGCAACGACGCAATTTTGTCTACCTCGAG 781  
721 CAAGAATGTGCTCTTCAATGTGCTTGTGTGCAACGACGCAATTTTGTCTACCTCGAG 780  
782 TCTCACAAGATGGCGTGGAGACCACTTTCAAGTGAATCATCTGGGGCACTTCTACCT 841  
781 TCTCACAAGATGGCGTGGAGACCACTTTCAAGTGAATCATCTGGGGCACTTCTACCT 840  
842 TGTCCAGCTCTCCAGATGTTTTGTGCGCTCAGCTCTCGCCCGTGTCAATTTGGTGTCTC 901  
841 TGTCCAGCTCTCCAGATGTTTTGTGCGCTCAGCTCTCGCCCGTGTCAATTTGGTGTCTC 900  
902 CTCAGAGTCCCATCGATTTACAGATATTAACGACTCTCTGGGAAAACGAGCTTCAGTCG 961  
901 CTCAGAGTCCCATCGATTTACAGATATTAACGACTCTCTGGGAAAACGAGCTTCAGTCG 960  
962 CCTCTCCGAAACAAAACGACTATTGGGGGATGCTGGCTTATAACAGGTCCAGCTCTG 1021  
961 CCTCTCTCCGAAACAAAACGACTATTGGGGGATGCTGGCTTATAACAGGTCCAGCTCTG 1020  
1022 CAACATCTCTTCTCCAAACGAGTGCACCGTGCCTCTCCCGACGGGGGTACGTCGAA 1081  
1021 CAACATCTCTTCTCCAAACGAGTGCACCGTGCCTCTCCCGACGGGGGTACGTCGAA 1080  
1082 CGCAGTGCATCTCGAAATATGATGACTCCAAATTCATCGCAGCTGTGGGTGTACAC 1141  
1081 CGCAGTGCATCTCGAAATATGATGACTCCAAATTCATCGCAGCTGTGGGTGTACAC 1140

1142 ACTGCTGTTTACCTTGGCGAGGCGCTTTTCAACAAGTCCATGCAACAGGAGCTGCAACCAC 1201  
1141 ACTGCTGTTTACCTTGGCGAGGCGCTTTTCAACAAGTCCATGCAACAGGAGCTGCAACCAC 1200  
1202 CGTGTACTGTGCTGTGCTCCAGAACTGGAGGCTTGGGAGGATGTACTTCAACAACATG 1261  
1201 CGTGTACTGTGCTGTGCTCCAGAACTGGAGGCTTGGGAGGATGTACTTCAACAACATG 1260  
1262 CTGCGCTGATGCTCTTACCCAGAACTCAGACGCAAGACGCGCCCGGACCTGTGGGC 1321  
1261 CTGCGCTGATGCTCTTACCCAGAACTCAGACGCAAGACGCGCCCGGACCTGTGGGC 1320  
1322 GCTCAGGAGAGCTCATCCAGAACTGTCAGCGCTTCCAGCTCCGGCTAAAGTGAGCTCAG 1381  
1321 GCTCAGGAGAGCTCATCCAGAACTGTCAGCGCTTGGCAGCAGCTCCGGCTAAAGTGAGCTCAG 1380  
1382 AGCGATGGGCAACACACACCGCCCTGTGTGTCTCCCTCACGCAAGTGCAGGGGTGGG 1441  
1381 AGCGATGGGCAACACACACCGCCCTGTGTGTCTCCCTCACGCAAGTGCAGGGGTGGG 1440  
1442 CCGCTTCCAAATGTCCTCCACACAGATCCGCAAGTAAAGGAAATAGAGCAGTCAC 1501  
1441 CCGCTTCCAAATGTCCTCCACACAGATCCGCAAGTAAAGGAAATAGAGCAGTCAC 1500  
1502 AACAGAGTGAATAATCTTAAGTACCAATGGAAGCAGGGAATTCCTGGGCTAAAGTATCA 1561  
1501 AACAGAGTGAATAATCTTAAGTACCAATGGAAGCAGGGAATTCCTGGGCTAAAGTATCA 1560  
1562 CTCTTCTGGGGCTGGGCTAGGCATAGTCTCTTTGCTTTCTGGTGGGCTGTTTGA 1621  
1561 CTCTTCTGGGGCTGGGCTAGGCATAGTCTCTTTGCTTTCTGGTGGGCTGTTTGA 1620  
1622 GTAAAACTGCTTGTGTGTGTAGGTTCCGATCTCTCCTGAGAGAACCCAGCAATTTCT 1681  
1621 GTAAAACTGCTTGTGTGTGTAGGTTCCGATCTCTCCTGAGAGAACCCAGCAATTTCT 1680  
1682 TTCTTTTACTGTTATAGAAATAGCTGAGGTCCCTCGTCCCATCCAGCTACCCACGCGC 1741  
1681 TTCTTTTACTGTTATAGAAATAGCTGAGGTCCCTCGTCCCATCCAGCTACCCACGCGC 1740  
1742 CACACTGCAGCGGGGCTGGGCTTTCTCTACTTAGGGAAGAAAAGCAAGTGTTCAT 1801  
1741 CACCACTGCAGCGGGGCTGGGCTTTCTCTACTTAGGGAAGAAAAGCAAGTGTTCAT 1800  
1802 GCTCCTTGTGCTGATTCAGAGAGATAATGTTTCAATCTCCTGACCAAGACTGAGCC 1861  
1801 GCTCCTTGTGCTGATTCAGAGAGATAATGTTTCAATCTCCTGACCAAGACTGAGCC 1860  
1862 AGCTTAGCAACTGCTGGGAGACAAATCTCAGAACTTGTCCCGACGAGTGAGGATGACA 1921  
1861 AGCTTAGCAACTGCTGGGAGACAAATCTCAGAACTTGTCCCGACGAGTGAGGATGACA 1920  
1922 GTGACACCCAGAGGAGTAGAATACGAGAACTACAGGTGGCAAGTACTTGTCTATAGA 1981  
1921 GTGACACCCAGAGGAGTAGAATACGAGAACTACAGGTGGCAAGTACTTGTCTATAGA 1980  
1982 CTCTTGTCTAATGCTATGCAAAAAATTTCTTAGAGATATTAACAAATTTTCAAATCAT 2041  
1981 CTCTTGTCTAATGCTATGCAAAAAATTTCTTAGAGATATTAACAAATTTTCAAATCAT 2040  
2042 TCCCTTAGATACCTTTGAAAGGACGAGGAGCGTATATATCTTAAAGAAATACAGAGATAT 2101  
2041 TCCCTTAGATACCTTTGAAAGGACGAGGAGCGTATATATCTTAAAGAAATACAGAGATAT 2100  
2102 TTTGGGGGCGAGAGATAAAACGTTAGTTAATCCCTTTTGTCTGTCATCACTCACTCTCAGT 2161  
2101 TTTGGGGGCGAGAGATAAAACGTTAGTTAATCCCTTTTGTCTGTCATCACTCACTCTCAGT 2160  
2162 TCTCTTGTCTTCACTGTATCTTAAACCTGCTGCTGTGCTCGCATCTTATGCTTATATA 2221  
2161 TCTCTTGTCTTCACTGTATCTTAAACCTGCTGCTGTGCTCGCATCTTATGCTTATATA 2220

QY 2222 AAGAACATGCTTGAATATCAAAAAAAAAAAAAA 2256  
DB 2221 AAGAACATGCTTGAATATCAAAAAAAAAAAAAA 2255

RESULT 4  
LOCUS BD156478 2256 bp DNA linear PAT 17-JAN-2003  
DEFINITION Primer for synthesizing full-length cDNA and use thereof.  
ACCESSION BD156478  
VERSION BD156478.1 GI:27862236  
KEYWORDS JP 2002191363-A/11321.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens

REFERENCE  
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1. (bases 1 to 2256)  
Ota,T., Isogai,T., Nishikawa,T., Hayashi,K., Saito,K., Yamamoto,J.,  
Ishii,S., Sugiyama,T., Wakamatsu,A., Nagai,K. and Otsuki,T.  
TITLE Primer for synthesizing full-length cDNA and use thereof  
JOURNAL Patent: JP 2002191363-A 11321 09-JUL-2002;  
HELIIX RESEARCH INSTITUTE

COMMENT  
OS Homo sapiens (human)  
PN JP 2002191363-A/11321  
PD 09-JUL-2002  
PF 28-JUL-2000 JP 2000280990  
PI TOSHIO OTA, TAKAO ISOGAI, TETSUO NISHIKAWA, KOJI HAYASHI, KAORU  
PI SAITO,  
PI JUNICHI YAMAMOTO, SHIZUKO ISHII, TOMOYASU SUGIYAMA, AI WAKAMATSU,  
PI KEIICHI NAGAI, TETSUO OTSUKI  
PC

C12N15/09, C07K14/47, C07K16/18, C12N1/15, C12N1/19, C12N1/21, C12N5/ PC  
10,  
PC, C12P21/02, C12Q1/68//C12P21/09, G06F17/30, C12N15/00, C12N5/00 CC  
Primer for synthesizing full-length cDNA and use thereof FH Key

FT CDS Location/Qualifiers  
1..2256  
1. .2256  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"

FEATURES  
source

ORIGIN

Query Match 99.5%; Score 2253.4; DB 6; Length 2256;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 2254; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 CAGTGCAGCGCTGAGCGGTGCGGCCCGCGAGCGCGCGGTCTCGTTGGAGCGGAGT 61  
DB 1 CAGTGCAGCGGTGAGCGGTGCGGCCCGCGAGCGCGCGGTCTCGTTGGAGCGGAGT 60

QY 62 GAGTTCCTGAGCGAGTGGACCGCGGCGAGCGGCGATAGGGGGCGAGTCCCTCCACAGTC 121  
DB 61 GAGTTCCTGAGCGAGTGGACCGCGGCGGCGGATAGGGGGCGAGTCCCTCCACAGTC 120

QY 122 AGCCATGGCAGCGCTGCGCTAGCGGGGCTGGACGACGACGAGTGGAGCGAGTGGCC 181  
DB 121 AGCCATGGCAGCGCTGCGCTAGCGGGGCTGGACGACGAGTGGAGCGAGTGGCC 180

QY 182 TCCGGCTGGAGGAGAGAACCAACGAGCGGCTGGTTTACTAGCCAAATCAACCGGA 241  
DB 181 TCCGGCTGGAGGAGAGAACCAACGAGCGGCTGGTTTACTAGCCAAATCAACCGGA 240

QY 242 GAGGAGACTCAGTGGGACATCCAAATCTGAAAAAGAAACGAGTGGCGAGGATTT 301  
DB 241 GAGGAGACTCAGTGGGACATCCAAATCTGAAAAAGAAACGAGTGGCGAGGATTT 300

QY 302 GCCATACGATGGGAACAAGAACTGATGAGAAACGACAAAGTGTGTTTGTGACCATAT 361  
DB 301 GCCATACGATGGGAACAAGAACTGATGAGAAACGACAAAGTGTGTTTGTGACCATAT 360

QY 362 AAATAAAGAACCACTACTTGGACCAAGACTGGCGTTTACTGTGGATGATAATCCGAC 421

DB 361 AAATAAAGAACCACTACTTGGACCAAGACTGGCGTTTACTGTGGATGATAATCCGAC 420  
QY 422 CAAGCCAAACCCCGGCAAGATACGACGCGAGCACCTGCGCATGGAAATTTCTCCAGGG 481  
DB 421 CAAGCCAAACCCCGGCAAGATACGACGCGAGCACCTGCGCATGGAAATTTCTCCAGGG 480

QY 482 CCGGATTTCACTGGCAAAAGTGTGTGCTACTCGAGCTAATTCAGGAATAGGGTTGGA 541  
DB 481 CCGGATTTCACTGGCAAAAGTGTGTGCTACTCGAGCTAATTCAGGAATAGGGTTGGA 540

QY 542 AACCCGCCAAGTCTTTTGGCCCTCCATGGTGCACATGTGATCTTTGGCCTGCAGGAACATGGC 601  
DB 541 AACCCGCCAAGTCTTTTGGCCCTCCATGGTGCACATGTGATCTTTGGCCTGCAGGAACATGGC 600

QY 602 AAGGCGGAGTGAAGCAGTGTACGGATTTTGAAGAAAGGCAATAGGCAAGTAGAAGC 661  
DB 601 AAGGCGGAGTGAAGCAGTGTACGGATTTTGAAGAAAGGCAATAGGCAAGTAGAAGC 660

QY 662 AATGACCCCTGGACCTCGCTCTCGCTAGCGTGCAGCATTTTGTGGAAGCATTTCAAGGC 721  
DB 661 AATGACCCCTGGACCTCGCTCTCGCTAGCGTGCAGCATTTTGTGGAAGCATTTCAAGGC 720

QY 722 CAAGAATGTGCTCTTCAATGTCTTGTGTGCAACGACGCAACTTTTGTCTTACCCCTGGAG 781  
DB 721 CAAGAATGTGCTCTTCAATGTCTTGTGTGCAACGACGCAACTTTTGTCTTACCCCTGGAG 780

QY 782 TCTCACCAGAGATGGCTGGAGACACCTTTCAAGTGAATCATCTGGGCACTTTTACCT 841  
DB 781 TCTCACCAGAGATGGCTGGAGACACCTTTCAAGTGAATCATCTGGGCACTTTTACCT 840

QY 842 TGTCCAGCTCTCCAGGATGTTTGTGCGCTCAGCTCTCTGCCCTGTCTATTGTGGTCTC 901  
DB 841 TGTCCAGCTCTCTCCAGGATGTTTGTGCGCTCAGCTCTCTGCCCTGTCTATTGTGGTCTC 900

QY 902 CTCAGAGTCCCATCGATTTACAGATATTAAAGCATCTCTTTGGGAAAACCTGAGCTCG 961  
DB 901 CTCAGAGTCCCATCGATTTACAGATATTAAAGCATCTCTTTGGGAAAACCTGAGCTCG 960

QY 962 CCTCTCTCCACAAAAGAGACTATTTGGCGATGTGGCTTATAACAGTCCCAAGTCTG 1021  
DB 961 CCTCTCTCCACAAAAGAGACTATTTGGCGATGTGGCTTATAACAGTCCCAAGTCTG 1020

QY 1022 CAACATCTCTCTCCAAAGAGCTGACCGTCCCTCTCCCAACGCGGGGTCACTCGGAA 1081  
DB 1021 CAACATCTCTCTCCAAAGAGCTGACCGTCCCTCTCCCAACGCGGGGTCACTCGGAA 1080

QY 1082 CGCAGTGCATCTCGGAAATATGATGTACTCCAAATTCATCGCAGCTGGTGGGTGTACAC 1141  
DB 1081 CGCAGTGCATCTCGGAAATATGATGTACTCCAAATTCATCGCAGCTGGTGGGTGTACAC 1140

QY 1142 ACTGTCTGTCTTGGCGGCGCTTTCACCAAGTCCATGCAACAGGAGCTGCCACCAAC 1201  
DB 1141 ACTGTCTGTCTTGGCGGCGCTTTCACCAAGTCCATGCAACAGGAGCTGCCACCAAC 1200

QY 1202 CGTGTACTGTGTCTGCCAGAACTGGAGGGTCTGGAGGGATGTACTTCAACAACTG 1261  
DB 1201 CGTGTACTGTGTCTGCCAGAACTGGAGGGTCTGGAGGGATGTACTTCAACAACTG 1260

QY 1262 TTGCGCTGCATGCGCTTCAACAGAGCTCAGAGCGAAGAGAGCGGCCCGGAACCTGTGGGC 1321  
DB 1261 TTGCGCTGCATGCGCTTCAACAGAGCTCAGAGCGAAGAGAGCGGCCCGGAACCTGTGGGC 1320

QY 1322 GCTCAGCGAGAGGCTGATCCCAAGACCGCTTTGGCAGCCAGTCCCGCTAAGTGGAGCTCAG 1381  
DB 1321 GCTCAGCGAGAGGCTGATCCCAAGACCGCTTTGGCAGCCAGTCCCGCTAAGTGGAGCTCAG 1380

QY 1382 AGCGATGGGACACACACCGCCCTGTGTGTGTCTCCCTCAGCAAGTCCAGGGCTGGG 1441  
DB 1381 AGCGATGGGACACACACCGCCCTGTGTGTGTCTCCCTCAGCAAGTCCAGGGCTGGG 1440

QY 1442 CCCCTTCCAAATGTCCCTCCAAACAGATCCGCAAGATTAAGGAATAAGAGCAGTCAAC 1501



Db 1441 CCCTTCCAAATGTCCTCCAAACACAGATCCGCAAGAGTAAGAAATAGAGCAATCAC 1500

QY 1502 AACAGAGTGAATAATCTTAAGTACCAATGGAAGCAGGGAATTCCTGGGTAAGATATCA 1561

Db 1501 AACAGAGTGAATAATCTTAAGTACCAATGGAAGCAGGGAATTCCTGGGTAAGATATCA 1560

QY 1562 CTTTCTGGGCTGGGCTAGGCATAGGTCTCTTTGCTTTCTGTGTGTGGCTGTTGAAA 1621

Db 1561 CTTTCTGGGCTGGGCTAGGCATAGGTCTCTTTGCTTTCTGTGTGTGGCTGTTGAAA 1620

QY 1622 GTAAAAACCTGCTGTGTGTAGTTCGGTATCTCCCTGAGAGACACCAAGCAATCTCT 1681

Db 1621 GTAAAAACCTGCTGTGTGTAGTTCGGTATCTCCCTGAGAGACACCAAGCAATCTCT 1680

QY 1682 TTCTTTTACTGTTATAGATAGCTGAGTCCCTCTGTCCTCCATCCAGTACCAACCAAGC 1741

Db 1681 TTCTTTTACTGTTATAGATAGCTGAGTCCCTCTGTCCTCCATCCAGTACCAACCAAGC 1740

QY 1742 CACCACTGACGCGGGCTGGCTTCTCTTACTTAGGGAAGAAAAAGCAAGTGTTCAT 1801

Db 1741 CACCACTGACGCGGGCTGGCTTCTCTTACTTAGGGAAGAAAAAGCAAGTGTTCAT 1800

QY 1802 GCTCCTTGTGCTGATTCAGAGAGATATTTGTTTCAATCTGACCAAGCAAGTGTTCAT 1861

Db 1801 GCTCCTTGTGCTGATTCAGAGAGATATTTGTTTCAATCTGACCAAGCAAGTGTTCAT 1860

QY 1862 AGCTTAGCAACTGCTGGGAGAGACAAATCTCAGAACCTTGTCCAGCCAGTGAGGATGACA 1921

Db 1861 AGCTTAGCAACTGCTGGGAGAGACAAATCTCAGAACCTTGTCCAGCCAGTGAGGATGACA 1920

QY 1922 GTGACACCCAGAGGAGTAGAATACGAGAACTACCAAGTGGCAAGTACTTGTCTATAGA 1981

Db 1921 GTGACACCCAGAGGAGTAGAATACGAGAACTACCAAGTGGCAAGTACTTGTCTATAGA 1980

QY 1982 CTCCTTTGCTTAATGCTATGCAAAAATCTCTTAGAGATTTATACAAATTTTCAATCAT 2041

Db 1981 CTCCTTTGCTTAATGCTATGCAAAAATCTCTTAGAGATTTATACAAATTTTCAATCAT 2040

QY 2042 TCCTTAGATACCTTGAAGGCGAGGAGGAGCGTATATATCTTAAGAAATACACAGGATAT 2101

Db 2041 TCCTTAGATACCTTGAAGGCGAGGAGGAGCGTATATATCTTAAGAAATACACAGGATAT 2100

QY 2102 TTTGGGGGCGAGAGATAAAGCTTATAGTTAATCCCTTTGCTGTCTCAATCACAGTCTCAGT 2161

Db 2101 TTTGGGGGCGAGAGATAAAGCTTATAGTTAATCCCTTTGCTGTCTCAATCACAGTCTCAGT 2160

QY 2162 TCTCTTGTCTTCACTGATTTAAACCTCTGCTGTGCTCCATCTTCTTATATAA 2221

Db 2161 TCTCTTGTCTTCACTGATTTAAACCTCTGCTGTGCTCCATCTTCTTATATAA 2220

QY 2222 AAGAACATGCTTGAATATCAAAAAAAGAAAAA 2256

Db 2221 AAGAACATGCTTGAATATCAAAAAAAGAAAAA 2255

# RESULT 5 AK027626

LOCUS AK027626 2256 bp mRNA linear PRI 01-AUG-2002

DEFINITION Homo sapiens cDNA FLJ14720 fis, clone NT2RP3001495, highly similar to Human oxidoreductase (HHCMA56) mRNA.

ACCESSION AK027626

VERSION AK027626.1 GI:14042433

KEYWORDS oligo capping; fis (full insert sequence).

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

## REFERENCE AUTHORS

1  
Isogai, T., Ota, T., Hayashi, K., Sugiyama, T., Otsuki, T., Suzuki, Y., Nishikawa, T., Nagai, K., Sugano, S., Shiratori, A., Sudo, H., Wagaesuma, M., Hosoliri, T., Kaku, Y., Kodaira, H., Kondo, H., Sugawara, M., Takanashi, M., Chiba, Y., Ishida, S., Murakawa, K., Ono, Y., Takiguchi, S., Watanabe, S., Kimura, K., Murakami, K.,

Ishii, S., Kawai, Y., Saito, K., Yamamoto, J., Wakamatsu, A., Nakamura, Y., Nagahara, K., Masuho, Y., Ninomiya, K. and Iwayanagi, T.  
NEDO human cDNA sequencing project  
Unpublished  
2 (bases 1 to 2256)  
Isogai, T. and Otsuki, T.  
Direct Submission  
Submitted (10-MAY-2001) Takao Isogai, Helix Research Institute, Genomics Laboratory; 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan (E-mail: genomics@hri.co.jp, Tel: 81-438-52-3975, Fax: 81-438-52-3986)  
NEDO human cDNA sequencing project supported by Ministry of Economy, Trade and Industry of Japan; cDNA full insert sequencing: Research Association for Biotechnology; cDNA library construction: 5'- and 3'-end one pass sequencing and clone selection: Helix Research Institute (supported by Japan Key Technology Center etc.) and Department of Virology, Institute of Medical Science, University of Tokyo.  
Location/Qualifiers  
1. .2256  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="NT2RP3001495"  
/cell\_line="NT2"  
/cell\_type="teratocarcinoma"  
/clone\_lib="NT2RP3"  
/notes="cloning vector: pME18SFL3-mRNA from NT2 neuronal precursor cells after 2-weeks retinoic acid (RA) induction."

## ORIGIN

Query Match 99.5%; Score 2253.4; DB 9; Length 2256;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 2254; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 CAGTGGCGAGCGGTGAGCGGTGCGGCCCGGACCGCGCGCGGTCTCGTTGGAGCGGAGT 61

Db 1 CAGTGGCGAGCGGTGAGCGGTGCGGCCCGGACCGCGCGCGGTCTCGTTGGAGCGGAGT 60

QY 62 GAGTTCTGACGCGAGTGGAACCGGCGAGCGGCGGATAGGGGGCGAGTGCCTCCAGTCC 121

Db 61 GAGTTCTGACGCGAGTGGAACCGGCGAGCGGCGGATAGGGGGCGAGTGCCTCCAGTCC 120

QY 122 AGCCATGGCAGCGCTCGGTACGCGGGGCTGGACGACGACGACGAGTGGAGCGAGTGC 181

Db 121 AGCCATGGCAGCGCTCGGTACGCGGGGCTGGACGACGACGAGTGGAGCGAGTGC 180

QY 182 TCCGGGCTGGAGGAGAGAAACCAACCAAGGACGCGTGGTTTACTAGCCCAATCACACCGA 241

Db 181 TCCGGGCTGGAGGAGAGAAACCAACCAAGGACGCGTGGTTTACTAGCCCAATCACACCGA 240

QY 242 GGAGAGAGCTCAGTGGGAAACATCCAAAACTGGAAGAAAGAAAGAGTGGCAGGAGATT 301

Db 241 GGAGAGAGCTCAGTGGGAAACATCCAAAACTGGAAGAAAGAAAGAGTGGCAGGAGATT 300

QY 302 GCCATACGATGGAAACAAAGAACTGATGAGAACGACCAAGTGTCTTTCTTGACCATAT 361

Db 301 GCCATACGATGGAAACAAAGAACTGATGAGAACGACCAAGTGTCTTTCTTGACCATAT 360

QY 362 AATAAAGAACCACTTCTTGGACCAAGACTGGCGTTTACTGTGATGATTAATCCGAC 421

Db 361 AATAAAGAACCACTTCTTGGACCAAGACTGGCGTTTACTGTGATGATTAATCCGAC 420

QY 422 CAGCCCAACCCCGGCAAGATACGACGCGCAGCACCACTGCCATGGAAATTTCCAGGG 481

Db 421 CAGCCCAACCCCGGCAAGATACGACGCGCAGCACCACTGCCATGGAAATTTCCAGGG 480

QY 482 CCGGGATTTCACTGGCAAGTGGTGTGTCCTGAGCTTAATTCAGGAATAGGGTTGGA 541

Db 481 CCGGGATTTCACTGGCAAGTGGTGTGTCCTGAGCTTAATTCAGGAATAGGGTTGGA 540

QY 542 AACCCGCAAGTCTTTTGGCCCTCCATGGTGCACATGTGATCTTGGCCCTGCGAGCAATGGC 601



/chromosome="16"  
/map="16q23.2"  
/note="FRA16D"  
105..1349  
/codon\_start=1  
/product="FOR II protein"  
/protein\_id="AA082054.1"  
/db\_xref="GI:8927391"  
/translation="MAALRYAGLDDTDEBELPGWEERTTKDGVVYVYAHTEKTCW  
EHPKTKRKVAGDLPVGEQETDENGQVFFVDHINKRTYLDPLAFVDDNPFTKPT  
TRQYDSTTAMEILOQRDTGKVVVTGANSIGFTAKSPALHGAHVILACRNMAR  
ASRAVLEEBHEWVHETVLDLALLRSVQHFABFAKKNVPVHLVLCNAATPALBP  
SLTKDGETTFQVNHLEFVLIQDLVLCRSAPRVIVVSSSHRETFDINDSLKLD  
FSLSKNDYMWAMLYNRSLKLNILFSLHRLSPRGVTSNAVHPGNMYSNIHRS  
WVVTLLFLARPPTKSMQOGAATTVYCAAVPELEGLGMYFNCCRCMPSPEAQSEE  
TARTLWALSERLLIQERLGSQS"

ORIGIN

Query Match	97.8%; Score 2213.4; DB 9; Length 2219;
Best Local Similarity	99.8%; Pred. No. 0;
Matches 2214; Conservative 2; Mismatches 3; Indels 0; Gaps 0;	
QY 22	TCGGGCCCGACGCGCGGGTCTCGTTTGGAGCGGAGTGAGTTCCTGACGAGTGAC 81
DB 1	TCGGGCCCGACGCGCGGGTCTCGTTTGGAGCGGAGTGAGTTCCTGACGAGTGAC 60
QY 82	CCGSCAGCGGCGATAGGGGGCGAGTGCCTCCACAGTCAGCCATGGCAGCGTGGCT 141
DB 61	CCGSCAGCGGCGATAGGGGGCGAGTGCCTCCACAGTCAGCCATGGCAGCGTGGCT 120
QY 142	ACGCGGGCTGGACACACGACAGTGGAGGACGAGTGCCTCCGGGGTGGGAGGAGAA 201
DB 121	ACGCGGGCTGGACACACGACAGTGGAGGACGAGTGCCTCCGGGGTGGGAGGAGAA 180
QY 202	CCACCAAGAGCGGTGGTTTACGCGCAATCAACGAGGAGAGACTCAGTGGGAAC 261
DB 181	CCACCAAGAGCGGTGGTTTACGCGCAATCAACGAGGAGAGACTCAGTGGGAAC 240
QY 262	ATCCAAAACCTGGAAAAGAAAAGAGTGGCAGGAGATTTGCCATACGATGGGAAC 321
DB 241	ATCCAAAACCTGGAAAAGAAAAGAGTGGCAGGAGATTTGCCATACGATGGGAAC 300
QY 322	AAACTGATGAGAACCGCAAGTGTGTTTGTGACCATATAATAAAGAACCACTACT 381
DB 301	AAACTGATGAGAACCGCAAGTGTGTTTGTGACCATATAATAAAGAACCACTACT 360
QY 382	TGGACCAAGACTGGCGTTTACTGTGATGATTAATCCGACCAAGCCAAACCCGGCAA 441
DB 361	TGGACCAAGACTGGCGTTTACTGTGATGATTAATCCGACCAAGCCAAACCCGGCAA 420
QY 442	GATACGAGCGCAGCACCACTGCCATGAAATTTCTCCAGGGCGGGATTTCACTGGCAAAG 501
DB 421	GATACGAGCGCAGCACCACTGCCATGAAATTTCTCCAGGGCGGGATTTCACTGGCAAAG 480
QY 502	TGGTGTGGTCACTGGAGCTAAATTCAGGAAATAGGGTTTCAAAACCGCAAGTCTTTGCC 561
DB 481	TGGTGTGGTCACTGGAGCTAAATTCAGGAAATAGGGTTTCAAAACCGCAAGTCTTTGCC 540
QY 562	TCCATGTGACATGTGATCTTGGCTCGCAGGAACATGGCAAGGGCGAGTGAACAGTGT 621
DB 541	TCCATGTGACATGTGATCTTGGCTCGCAGGAACATGGCAAGGGCGAGTGAACAGTGT 600
QY 622	CACGCAATTTAGAAATGGCAATAGCAAGTAGAAGCAATCACCTGGACCTCGCTC 681
DB 601	CACGCAATTTAGAAATGGCAATAGCAAGTAGAAGCAATCACCTGGACCTCGCTC 660
QY 682	TGCTCCGTAGCGTGCAGCATTTTCTGAAAGCATTAAGGGCCAAAGATGTGCTTTCATG 741
DB 661	TGCTCCGTAGCGTGCAGCATTTTCTGAAAGCATTAAGGGCCAAAGATGTGCTTTCATG 720
QY 742	TGCTTGTGCAACGAGCAAACTTTTGTCTCTACCTGGAGTCTCACAAAGATGGCTGG 801
DB 721	TGCTTGTGCAACGAGCAAACTTTTGTCTCTACCTGGAGTCTCACAAAGATGGCTGG 780

QY 802	AGACCACTTTCAAGTGAATCATCTGGGACATTTCTACCTTGTCCAGTCTCTCCAGGATG 861
DB 781	AGACCACTTTCAAGTGAATCATCTGGGACATTTCTACCTTGTCCAGTCTCTCCAGGATG 840
QY 862	TTTTGTCCGCTCAGCTCTCTGCCGCTGTCATTGTGGTCTCCTCAGAGTCCCATGATTA 921
DB 841	TTTTGTCCGCTCAGCTCTCTGCCGCTGTCATTGTGGTCTCCTCAGAGTCCCATGATTA 900
QY 922	CAGATATTAAGACTCTCTTGGGAAAATCGACTTCAGTCCGCTCTCTCCACAAAACG 981
DB 901	CAGATATTAAGACTCTCTTGGGAAAATCGACTTCAGTCCGCTCTCTGCAAAAACG 960
QY 982	ACTATTGGGCGATCTCGCTTATTAACAGGTCCAAGCTCTGCAACATCTCTCTCCACG 1041
DB 961	ACTATTGGGCGATCTCGCTTATTAACAGGTCCAAGCTCTGCAACATCTCTCTCCACG 1020
QY 1042	AGTCGACCGTGGCTCTCCGCCACGCGGGGTACGTGAAACGAGTGATCTCTGGAATA 1101
DB 1021	AGTCGACCGTGGCTCTCCGCCACGCGGGGTACGTGAAACGAGTGATCTCTGGAATA 1080
QY 1102	TGATGTAATCCAACTTCATCGCAGCTGGTGGGTGACACACTGCTGTTTACCTTGGGA 1161
DB 1081	TGATGTAATCCAACTTCATCGCAGCTGGTGGGTGACACACTGCTGTTTACCTTGGGA 1140
QY 1162	GGCCTTTCAACCAAGTCCATGCAACAGGAGCTGCACACCGTGTACTGTGCTGTGCTCC 1221
DB 1141	GGCCTTTCAACCAAGTCCATGCAACAGGAGCTGCACACCGTGTACTGTGCTGTGCTCC 1200
QY 1222	CAGAACTGGAGGGTCTGGAGGGATGATCTTCAAACAACGTGCTGCGGTGCATGCGCTCAC 1281
DB 1201	CAGAACTGGAGGGTCTGGAGGGATGATCTTCAAACAACGTGCTGCGGTGCATGCGCTCAC 1260
QY 1282	CAGAACTGAGGAGAGAGACGCGGACCTGTGGGGCTCAGGAGAGGCTGTATCC 1341
DB 1261	CAGAACTGAGGAGAGAGACGCGGACCTGTGGGGCTCAGGAGAGGCTGTATCC 1320
QY 1342	AAGAACTGGCAGCAGTCCGGCTTAAAGTGAAGTCTCAGAGCGATGGGACACACAC 1401
DB 1321	AAGAACTGGCAGCAGTCCGGCTTAAAGTGAAGTCTCAGAGCGATGGGACACACAC 1380
QY 1402	CGCCTGTGTGTGCTCCCTCAGCAAGTGCAGGGCTGGGCGCTTCAAATGCTCCCTCC 1461
DB 1381	CGCCTGTGTGTGCTCCCTCAGCAAGTGCAGGGCTGGGCGCTTCAAATGCTCCCTCC 1440
QY 1462	AACACAGATCGCAAGAGTAAAGAAATAGAGCAGTCAACACAGAGTGAATAATCTTAA 1521
DB 1441	AACACAGATCGCAAGAGTAAAGAAATAGAGCAGTCAACACAGAGTGAATAATCTTAA 1500
QY 1522	GTACCAATGGAAAGCAGGAAATCTCGGGTAAAGTATCACTTTCTGGGGCTGGGCTAG 1581
DB 1501	GTACCAATGGAAAGCAGGAAATCTCGGGTAAAGTATCACTTTCTGGGGCTGGGCTAG 1560
QY 1582	GCAATGCTCTTTGCTTTCTGTGTGGCTGTTGAAAGTAAACCTGCTTGTGTGTG 1641
DB 1561	GCAATGCTCTTTGCTTTCTGTGTGGCTGTTGAAAGTAAACCTGCTTGTGTGTG 1620
QY 1642	TAGTTCCGTATCTCCCTGAGAGACACAGCAATTTCTCTTTCTTTTACTGTTATAGAT 1701
DB 1621	TAGTTCCGTATCTCCCTGAGAGACACAGCAATTTCTCTTTCTTTTACTGTTATAGAT 1680
QY 1702	AGCTGAGGTCCCTCTGCTCCATCCAGTACCACACGCGCACCTGACGCGGGGCT 1761
DB 1681	AGCTGAGGTCCCTCTGCTCCATCCAGTACCACACGCGCACCTGACGCGGGGCT 1740
QY 1762	GGCCTTCTCTACTTAGGGGAGAAAAGCAAGTGTTCACCTGCTCTTGTGCTATGATCC 1821
DB 1741	GGCCTTCTCTACTTAGGGGAGAAAAGCAAGTGTTCACCTGCTCTTGTGCTATGATCC 1800
QY 1822	AGGAGATAATTTGTTTCTATCTGACCAAGACTGAGCCAGCTTAGCAACTGCTGGGGA 1881
DB 1801	AGGAGATAATTTGTTTCTATCTGACCAAGACTGAGCCAGCTTAGCAACTGCTGGGGA 1860

```
QY 1882 GACAAATCTCAGAACCTTGTCCAGCCAGTGGAGATGACAGTGAACACCAGAGGAGTAG 1941
Db 1861 GACAAATCTCAGAACCTTGTCCAGCCAGTGGAGATGACAGTGAACACCAGAGGAGTAG 1920
QY 1942 AATACGAGAACTACAGGTCGCAAGTACTTGTCTATAGACTCCCTTGTCTAATGCTATGC 2001
Db 1921 AATACGAGAACTACAGGTCGCAAGTACTTGTCTATAGACTCCCTTGTCTAATGCTATGC 1980
QY 2002 AAAAAATCTTTAGAGATTATAACAAATTTTTTCAAAATCATTTCTTAGATACCTTGAAAGG 2061
Db 1981 AAAAAATCTTTAGAGATTATAACAAATTTTTTCAAAATCATTTCTTAGATACCTTGAAAGG 2040
QY 2062 CAGGAAGGAGCGGTATATATCTTAAGATATACACAGATATTTTGGGGGCGCAGAGATAAA 2121
Db 2041 CAGGAAGGAGCGGTATATATCTTAAGATATACACAGATATTTTGGGGGCGCAGAGATAAA 2100
QY 2122 ACGTTAGTTAATCCCTTGTCTGTCTCAATACAGTCTCAGTCTCTTGTCTTACATTTGTA 2181
Db 2101 ACGTTAGTTAATCCCTTGTCTGTCTCAATACAGTCTCAGTCTCTTGTCTTACATTTGTA 2160
QY 2182 CTTAAACCTCTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGT 2240
Db 2161 CTTAAACCTCTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGT 2219

RESULT 7
HSU13395
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

HSU13395
Human oxidoreductase (HHCMA56) mRNA, complete cds.
U13395.1
GI:538131
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES
source

1..1475
/organism="Homo sapiens"
/mol_type="mRNA"
/isolate="two Year old female"
/db_xref="taxon:9606"
/chromosome="16"
/sex="female"
/tissue_type="hippocampus"
/clone_lib="hippocampus library, Stratagene catalog number
936205"
/dev_stage="juvenile"
/notes="DSEG number: D168432E"
1..1475
/gene="HHCMA56"
19..1131
/gene="HHCMA56"
/codon_start=1
/product="oxidoreductase"
/protein_id="AAA21465.1"
/db_xref="GI:538132"
translation="MARASEAVSRILEEWHKAEAMTLDIALRSVQHFAEPAKKN
VPLHVCNAAATFALPWSLTGKLETFQVNLGHFVLVOLLPGMFCQAQLLPVSLWS
POSPLDQLLTWENWTSVASLOKTTIGRCWLTITGPSSATSSPSCTVASPTRGH
VERSDRWSKVDVLOHSQVHTAVYLGEAFHQVHATGSCHHRVLCCEPRTGSRRD
VLQQLLPLHALTRSRERDGPDPVGLSERLIQERLAASPAKWSERMGMTHTRPVCVPS
```

```
ORIGIN
Query Match
Best Local Similarity 60.6%; Score 1372.4; DB 9; Length 1475;
Matches 1459; Conservative 0; Mismatches 11; Indels 10; Gaps 7;
RKCAGPLNPVPTQIRKSGKNGSIHNRKVLKYQWEAGNSWGKVSLEFWGWARHSLC
PLWVACLKVKTWLACRPRISLEKHQFSSFCYRIA"
QY 579 ATCTTGGCTCTCAGGAACATCGCAAGCGCGAGTGAAGCAGTGTCTACGCATTTTGAAGAA 638
Db 1 ATCTTGGCTCTCAGGAACATCGCAAGCGCGAGTGAAGCAGTGTCTACGCATTTTGAAGAA 60
QY 639 TGGCATAAAGCAAGGTAGAAGCAATGACCCCTGGACCTCGCTCTCTCCGTAGCGTGCAG 698
Db 61 TGGCATAAAGCAAGGTAGAAGCAATGACCCCTGGACCTCGCTCTCTCCGTAGCGTGCAG 120
QY 699 CATTTTGTGAAGCATTTCAAGGCCAAGATGTGCTCTTTCATGTGCTTGTGTGCAACGCA 758
Db 121 CATTTTGTGAAGCATTTCAAGGCCAAGATGTGCTCTTTCATGTGCTTGTGTGCAACGCA 180
QY 759 GCAACTTTTGTCTTACCCCTCGAGTCTCACAAAGATGGCTGGAGACCAACCTTTCAAAGTG 818
Db 181 GCAACTTTTGTCTTACCCCTCGAGTCTCACAAAGATGGCTGGAGACCAACCTTTCAAAGTG 240
QY 819 AATCATCTGGGGCACTTCTACCTTGTCCAGTCTCCTCC--AGGATGTTTGTGCCGCTCAG 876
Db 241 AATCATCTGGGGCACTTCTACCTTGTCCAGTCTCCTCCAGGGATGTTTGTGCCGCTCAG 300
QY 877 CTCCTGCCCGTGTCTTGTGCTCTCCTCAGAGTCCATCGGATTTACAGATATTACGACT 936
Db 301 CTCCTGCCCGTGTCTTGTGCTCTCCTCAGAGTCCATCGGATTTACAGATATTACGACT 360
QY 937 CTTTGGGAAAACTGACCTTCACTGCTCGCTCTCTCCAAACAAAAACGACTATTGGGCGATGC 996
Db 361 CTTTGGGAAAACTGACCTTCACTGCTCGCTCTCTCCAAACAAAAACGACTATTGGGCGATGC 420
QY 997 TGGCTTTAAACAGTCTCAAGCTCTGCAACATCTCTTCTCCAAACGAGTGCACCGTCCG 1056
Db 421 TGGCTTTAAACAGTCTCAAGCTCTGCAACATCTCTTCTCCAAACGAGTGCACCGTCCG 480
QY 1057 TCTCCCAACGCGGGGTCACTGCGAAGCGAGT--CATCTGGAAATATGATGACTCTCA 1113
Db 481 TCT-CCCAACGCGGGGTCACTGCGAAGCGAGTGCATCTCGGAAATATGATGACTCTCA 539
QY 1114 ACATTTCATCGCAGCTGGTGGGTGTACACACTGCTGTTTACCTTGGCAGAGGCTTTTACCA 1173
Db 540 ACATTTCATCGCAGCTGGTGGGTGTACACACTGCTGTTTACCTTGGCAGAGGCTTTTACCA 599
QY 1174 AGTCCATGCAACAGGAGCTGCCACCCCGTGTACTGTGCTGTCTCCAGAACTGGAGG 1233
Db 600 AGTCCATGCAACAGGAGCTGCCACCCCGTGTACTGTGCTGTCTCCAGAACTGGAGG 659
QY 1234 GTCTGGGAGGATGTACTTCAACAACTGCTGCCGCTCATGCCCTCACCAGAACTCAGA 1293
Db 560 GTCTAGAGGAGATGTACTTCAACAACTGCTGCCGCTCATGCCCTCACCAGAACTCAGA 719
QY 1294 GCGAAGAGACGCGCCCGGACCTGTGGGCGCTCAGCGAGGCTGATCCAAAGAGCGGTG 1353
Db 720 GCGAAGAGACGCGCCCGGACCTGTGGGCGCTCAGCGAGGCTGATCCAAAGAGCGGTG 777
QY 1354 GCAGCGAGTCCGGCTAAGTGAGCTCAGAGCGGATGGGCAACACCCGCCCTGTGTGT 1413
Db 778 GCAGCGAGTCCGGCTAAGTGAGCTCAGAGCGGATGGGCAACACCCGCCCTGTGTGT 837
QY 1414 GTCCCTCTCAGCGAGTCCCGGGCTGGCGCTTCCAAATGTCCCTCCACACAGATCCG 1473
Db 838 GTCCCTCTCAGCGAGTCCCGGGCTGGCGCTTCCAAATGTCCCTCCACACAGATCCG 896
QY 1474 CAAGAGTAAAGGAAATAGAGCAGTCAACACAGAGTGAATAAATCTTAAGTACCAATGGGA 1533
Db 897 CAAGAGTAAAGGAAATAGAGCAGTCAACACAGAGTGAATAAATCTTAAGTACCAATGGGA 956
QY 1534 AGCAGGAAATTCCTGGGGTAAAGTATCACTTTTCTGGGGCTGGGCTAGGCATAGGTCTCT 1593
```

|||||  
957 AGCAGGAAATTCCTGGGTAAGATATCACATTTCTGGGGCTGGGCTAGGCATAGGTCTCT 1016  
|||||  
1594 TTGCTTTCTGGTGGCCCTGTTTGAAGTAAGAAACCTGCTTGGTGTGATAGGTTCGGTAT 1653  
|||||  
1017 TTGCTTTCTGGTGGCCCTGTTTGAAGTAAGAAACCTGCTTGGTGTGATAGGTTCGGTAT 1076  
|||||  
1654 CTCCTGGGAAGCAGCAGCAATCTCTTTCTTTTACTGTATAGAAATAGCCTGAGGTCC 1713  
|||||  
1077 CTCCTGGAGAGCAGCAGCAATCTCTTTCTTTTACTGTATAGAAATAGCCTGAGGTCC 1136  
|||||  
1714 CTTCTGCTCCATCCAGCTACCAACGCGCCACCTATGTCAGCGGGGGCTGCCCTTCTCCTA 1773  
|||||  
1137 CTTCTGCTCCATCCAGCTACCAACGCGCCACCTATGTCAGCGGGGGCTGCCCTTCTCCTA 1195  
|||||  
1774 CTTAGGGAAGAAAGCAAGTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1833  
|||||  
1196 CTTAGGGAAGAAAGCAAGTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1255  
|||||  
1834 TTTCTATCTCTCTGACCAAGCTGAGCAGCTTAGCAATCTAGCAATCTGCGGAGACAAATCTCAG 1893  
|||||  
1256 TTTCTATCTCTCTGACCAAGCTGAGCAGCTTAGCAATCTGCGGAGACAAATCTCAG 1315  
|||||  
1894 AACCTTGTCCAGCAGTGTAGGATGACAGTGACACCCAGAGGGAGTGAATACGCAAGAC 1953  
|||||  
1316 AACCTTGTCCAGCAGTGTAGGATGACAGTGACACCCAGAGGGAGTGAATACGCAAGAC 1375  
|||||  
1954 TACAGGTGCAAGTGTGTATAGATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2013  
|||||  
1376 TACAGGTGCAAGTGTGTATAGATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1435  
|||||  
2014 AGAGATTATACAAATTTTCAATCATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2053  
|||||  
1436 AGAGATTATACAAATTTTCAATCATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1475  
|||||

RESULT 8  
AF395123  
LOCUS  
DEFINITION Homo sapiens WWOXdelta6-8 mRNA, complete cds, alternatively spliced.  
ACCESSION AF395123  
VERSION AF395123.1 GI:15028462  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
REFERENCE 1 (bases 1 to 1732)  
AUTHORS Bednarek, A.K., Keck-Waggoner, C.L., Daniel, R.L., Lafin, K.J., Bergsagel, P.L., Kiguchi, K., Brenner, A.J. and Aldaz, C.M.  
TITLE WWOX, the FRA16D gene, behaves as a suppressor of tumor growth  
JOURNAL Cancer Res. 61 (22), 8068-8073 (2001)  
MEDLINE 21575822  
PUBMED 11719429  
REFERENCE 2 (bases 1 to 1732)  
AUTHORS Bednarek, A.K. and Aldaz, C.M.  
TITLE Direct Submission  
JOURNAL Submitted (25-JUN-2001) Carcinogenesis, The University of Texas, M.D. Anderson Cancer Center, Science Park-Research Division, Park Road 1C, Smithville, TX 78957, USA  
FEATURES  
source  
1..1732  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
134..838  
/note="aberrant WWOX transcript found in cancer cells"  
/codon\_start=1  
/evidence=not\_experimental  
/product="WWOXdelta6-8"  
/protein\_id="AAK81727.1"  
/db\_xref="GI:15028463"

/translation="MAALRYAGLDDTSDDELPPGWERTTKDGVVYVYAHNTBEKTQW  
EHPKTKRVRAGDLPVGEQETDQVFFVDHINKRTYLDPLRAFTVDDNPKPT  
THQRYDGSFTAMEILQGRDFTGVVGTGANSIGFETAKSFAHGHVILAENMAR  
ASEAVRILIEWOQGAATVYCAAVPELBLEGLGMFNNCCRCMPSPQAQSEETATLW  
AUSERLIQERLGSQSG"  
Query Match 51.9%; Score 1174; DB 9; Length 1732;  
Best Local Similarity 76.1%; Pred. No. 2.1e-310;  
Matches 1724; Conservative 0; Mismatches 0; Indels 540; Gaps 1;  
ORIGIN  
QY 1 GCAGTGGCAGGCGTGGAGCGTGGGCGCCGCGAGCGCGCGGGGCGAGTGGCTCCACAGT 60  
DB 9 GCAGTGGCAGGCGTGGAGCGTGGGCGCCGCGAGCGCGGGGCGAGTGGCTCCACAGT 68  
QY 61 TGAGTTCCTCAGCGAGTGGACCCGCGAGCGGGGCGATAGGGGGGCGAGTGGCTCCACAGT 120  
DB 69 TGAGTTCCTCAGCGAGTGGACCCGCGAGCGGGGCGATAGGGGGGCGAGTGGCTCCACAGT 128  
QY 121 CAGCCATGGCAGCGCTGCGCTACGCGGGGCTGGACGACACGAGCAGTGGAGCAGGCTGC 180  
DB 129 CAGCCATGGCAGCGCTGCGCTACGCGGGGCTGGACGACACGAGCAGTGGAGCAGGCTGC 188  
QY 181 CTCGGGCTGGGAGGAGAACCCAGGAGCAGCGCTGGGTTTACTACGCCAATCACACCG 240  
DB 189 CTCGGGCTGGGAGGAGAACCCAGGAGCAGCGCTGGGTTTACTACGCCAATCACACCG 248  
QY 241 AGAGAAAGTCTAGTGGGAAATCCAAAATCTGAAAAAAGAAAAAGAGTGGCAGGAGATT 300  
DB 249 AGAGAAAGTCTAGTGGGAAATCCAAAATCTGAAAAAAGAAAAAGAGTGGCAGGAGATT 308  
QY 301 TGGCATACGATGGGAAACAAAGAACTGATGAGAACGAGCAAGTGTGTTTGTGTGACCAT 360  
DB 309 TGGCATACGATGGGAAACAAAGAACTGATGAGAACGAGCAAGTGTGTTTGTGTGACCAT 368  
QY 361 TAAATAAAGAACCACTACTTGGACCCAGACTGGCGTTTACTGTGGATGATTAATCCGA 420  
DB 369 TAAATAAAGAACCACTACTTGGACCCAGACTGGCGTTTACTGTGGATGATTAATCCGA 428  
QY 421 CCAAGCCAAACCCCGCAAGATACGACGGGAGCACCCTGCGCATGGAATTTCCAGG 480  
DB 429 CCAAGCCAAACCCCGCAAGATACGACGGGAGCACCCTGCGCATGGAATTTCCAGG 488  
QY 481 GCGGGATTCTCTGCAAAAGTGTGTGTCTACTGGAGCTAAATTCAGGAATAGGGTTCG 540  
DB 489 GCGGGATTCTCTGCAAAAGTGTGTGTCTACTGGAGCTAAATTCAGGAATAGGGTTCG 548  
QY 541 AAACCCCAAGTCTTTTGGCCCTCCATGGTGCACATGTGATCTTGGCTGAGGAACATGG 600  
DB 549 AAACCCCAAGTCTTTTGGCCCTCCATGGTGCACATGTGATCTTGGCTGAGGAACATGG 608  
QY 601 CAAGGGGAGTGAAGCAGTGTCAAGCATTTTGAAGAAATGGCATAAAGCCAGTAGAAG 660  
DB 609 CAAGGGGAGTGAAGCAGTGTCAAGCATTTTGAAGAAATGGCATAAAGCCAGTAGAAG 668  
QY 661 CAATGACCCCTGGACCTGCTCTGCTCGTAGCGTGGAGCATTTTGTGGAAGCATTTCAAGG 720  
DB 649 ----- 648  
QY 721 CCAAGAAATGTCCTCTTCAATGTCTGTGTGTAACGAGCAACTTTTGTCTTACCTGGA 780  
DB 649 ----- 648  
QY 781 GTCTCACCAGAGTGGCCTGGAGACCCCTTTCAGTGAATCATCTGGGCACTTCTTACC 840  
DB 649 ----- 648  
QY 841 TTGTCCAGCTCCTCCAGGAGTGTGTCGCTCAGCTCCTGCGGTGTCATTTGTGTGTCT 900  
DB 649 ----- 648  
QY 901 CCTCAGAGTCCCATCGATTTTACAGATATTAAAGACTCTCTTGGGAAACTGGACTTCAGTC 960









```
/db_xref="taxon:10090"
/clone="MGC:25975 IMAGE:4241066"
/tissue_type="Kidney, normal. 5 month old male mouse."
/clone_lib="NCI CGAP_Kid14"
/lab_host="DH10B"
/note="Vector: pCMV-SPORT6"
1..2269
/gene="Wwox"
/note="synonym: WOX1"
/db_xref="LocusID:80707"
/db_xref="MGI:1931237"
77..11321
/codon_start=1
/product="WW-domain oxidoreductase"
/protein_id="AAH14716.1"
/db_xref="GI:15928476"
/db_xref="LocusID:80707"
/translation="MAALRYAGLDLDDSEDELPPGWEERTTKDGVVYANHTEKTQW
EHPKTKRRVAGDLPYGWEQEDENGQVFFVDHINKRTYLDPLRAFTVDNPKPT
TPQYDGSSTAMEILQGRDFTGKVLVTGANSIGIFETAKSFALGHVILACRLSR
ASEAVSRILEWHKAVEAMTLDLVLRSVQHFPAEFAKAKVSLHVLVGNAGTFALPW
GLTKDLETTFOVNHKGFLVQLLDVLCRSSPARVIVVSESHRFTDINDSSGKLD
LSRLSPRSDYAMLAYNSKLNILFSLNELHRLSPRGVTSNAVHPGNMYSAIHN
SWYKLLFLIARPFYSMQQGRATIVICAVAPELEGIGMYFNCCRLPSPERAQSE
TARAUWELSERLIQRLKSPSS"
misc_feature
128..217
/note="WW; Region: Domain with 2 conserved Trp (W)
residues"
/db_xref="CDD:smart00456"
misc_feature
437..1120
/note="FabG; Region: Dehydrogenases with different
specificities (related to short-chain alcohol
dehydrogenases) [secondary metabolites biosynthesis,
transport, and catabolism / General function prediction
only]"
/db_xref="CDD:COG1028"

ORIGIN

Query Match      48.4%; Score 1096.6; DB 10; Length 2269;
Best Local Similarity 72.6%; Pred. No. 3.9e-289;
Matches 1597; Conservative 0; Mismatches 534; Indels 70; Gaps 11;

QY 107 GGTGCTCCACAGTCAGCCATGCGAGCGTGCCTACGGGGCTGGACGACGACAG 166
Db 58 GGTACTTCTGCTCCAGCATGTCAGTCTGCGCTATGCGGGCTGGACGACGACAG 117
QY 167 TGAGGACGAGCTCCCTCGGGCTGGAGGAGAGAACACCAAGAGCGGTGGTACTA 226
Db 118 TGAGGATGAGCTCCCTCGGGCTGGAGGAGAGAACACCAAGAGCGGTGGTACTA 177
QY 227 CGCAATCACACGAGGAGAGACTCAGTGGGAAACATCCAAAACCTGAAAAGAAAACG 286
Db 178 CGCAATCACACTAGGAGAGAGACCAGTGGGAAACATCCGAAACCGGCAAGAGAAACG 237
QY 287 AGTGGCAGAGATTTCGCATACGATGGGNAAGAAGAACTGATGAGAACCGCAAGTGT 346
Db 238 GGTGCGAGGAGATTTCGCGTATGGATGGGAACAGGAGACTGACGAGAACCGCAAGTGT 297
QY 347 TTTTGTGACCATATAATAAAGAACCACTACTTGGACCCAGACTGGCGTTTACTGT 406
Db 298 TTTGTGTGACCATATAATAAAGAACTACTACTTGGATCCAAAGATTGGCAATTACCGT 357
QY 407 GATGATATCCGACCAAGCAACACCCCGGCAAGATACGCGGACGACCACTGCCAT 466
Db 358 GGACGATATCCGACGAGCCGACCAACAGACAGATACGCGGAGCAACCCGCGCAT 417
QY 467 GGAATTCACGGCGGGGATTTCATGCGCAAAGTGGTGTGGTCACTGGAGCTAATTC 526
Db 418 GGAGATCCTGCGAGGCGGGGACTTCACAGGCAAGTGGTCTGTTACTGGAGCAAAATC 477
QY 527 AGGAATAGGGTTCGAACCCCAAGCTTTTTCCTCCATGGTGACATGTGATCTTGGC 586
Db 478 GGAATAGGGTTCGAGACTGCCAAGTCTTTTGCCTCCATGGCGCGCATGTGATCTGGC 537
```

```
QY 587 CTGCAAGAACATGGCAAGGGCGAGTGAAGCAGTGTCAAGCAATTTTAGAAGAAATGGCAAT 646
Db 538 CTGTAGGAACCTGAGCAGAGCCAGCGAGGCCGTGTCGCGCATCTCGAGGAGTGGCACAA 597
QY 647 ACCAAGGTAGAAAGCAATGACCTCGACCTGCTCTGCTCGGTAGCTGAGCTGAGCAATTTGC 706
Db 598 AGCCAAGGTGGAAGCAATGACCTCGACCTGCGCGGTGCTCGGTAGCTGAGCTGAGCAATTTGC 657
QY 707 TCAAGCATTTCAAGGCCCAAGAATGCTCTTCAATGTGCTGTGTGTCGCAACGACGACAACTTT 766
Db 658 CGAAGCATTTCAAGGCCCAAGAATGTGCTCTTCAAGCTGTGTGTGTGTGTGTGTGTGTGTGT 717
QY 767 TGTCTTACCTCTGGAGTCTTCAACCAAGATGGCTGGAGACCACTTTTCAAGTGAATCATCT 826
Db 718 TGTCTTGTCTCTGGGCTTCAAGAGGATGGCTGGAGACCACTTTTCAAGTGAACCACT 777
QY 827 GGGGCACTTCTACCTTGTCCAGCTCTCCAGATGTTTGTGCGCTCAGCTCTGCTGCGCG 886
Db 778 GGGACATTTTCTACCTTGTCCAACTCTCTGAGGATGTTTGTGCGCTCAGCTCTGCGCGCG 837
QY 887 TGTCTATTGTGCTCTCTCAGAGTCCCATCGATTTACAGATATTAAAGTCTCTCTGGGAAA 946
Db 838 TGTCTATTGTGCTCTCTCAGAGTCCCATCGATTTACAGATATTAAAGTCTCTCTGGGAAA 897
QY 947 ACTGGACATTTCACTGCGCTCTCTCCAAACAAAACGACTATTGGGGGATGCTGGCTTTATAA 1006
Db 898 ACTTGAATCTCAGCGCTCTCTCTCCACCGAGGAGCGACTACTGGGCGCATCTGGCTTTATAA 957
QY 1007 CAGGTCCAAGCTCTGCAACATCTCTTCTCCAAAGAGCTGACCGTGGCTCTGCTGCGCG 1066
Db 958 CAGATCTAACTCTGCAATATCTCTTCTCCAAAGAGCTGACCGTGGCTCTGCTGCGCG 1017
QY 1067 CCGGGTCACTGCGAAGCGAGTGCATCTCTGGAATATGATGATCTCCAAATTCATCGCAG 1126
Db 1018 CCGGGTCACTTCACGCGAGTGCATCTCTGGGAACATGATGATCTCAGCCATCCATCGCAA 1077
QY 1127 CTGGTGGGTGTACACACTGCTGTTTACCTTGGCGAGCGCTTTTCCAAAGTCCATGCAACA 1186
Db 1078 CTCATGGGTGTACAAGTCTGTTTCACTTGGCGAGCGCTTCCAAAGTCCATGCAACA 1137
QY 1187 GGGAGTGCACACCGTGTACTGTCTGTCCAGAACTGGAGGGTCTGGAGGGTCTGGAGGGAT 1246
Db 1138 GGGAGTGTGACCACTGTCTACTGTGTCAGTCCCGCAGAGCTGGAGGGCTCGAGGGAT 1197
QY 1247 GTACTTTCAACAACTGTCGCGTGCATGCCCTCACCAAGAGCTCAGAGCGAAGAGACGGC 1306
Db 1198 GTATTTCAACAACTGTCGCGTGCATGCCCTCACCGAGGGTCTCAGAGTGAAGAGACAGC 1257
QY 1307 CCGGACCTCTGGGCGCTCAGCGAGAGGCTGATCCAAAGAACCGCTTGGCAGCCAGTCCGG 1366
Db 1258 TCGGGCGCTGTGGGAACTCAGTGAGAGGCTCATCCAGGACCGACTGGGTAGTCCATCCAG 1317
QY 1367 CTAAGTGGAGCTCAGAGCGGATGGGACACACACCGCGCTGTGTGTGTCCTCCCTCAGCA 1426
Db 1318 CTAATTTGGAACCCCAATATTTGGGCAAAACCCCACTTGTGTGTGTGTGTGTGTGTGTGTGT 1368
QY 1427 AGTGGCAGGGTGGGCGCTTCCAAATGTCCTCCAAACAGATCCGCAAGAGTGAAGCA 1486
Db 1369 CATCAGTACGCCCATCATCTGGGTGCCCTGAACTCCAGTACAGGCAATGAGAGTA 1428
QY 1487 AATAAGACAGTCAACAACAGAGTGAATAATCTTAAGTACCAATGGGAAGACGGGAATTC 1546
Db 1429 GATAAGAACAAATC-----AATGGAGAGAGGTATAACTAGTGGGAGACGGGAATTTTC 1481
QY 1547 TCGGGTAAAGTATACATTTTCTGGGCTGGGTAGGCATAGTCTCTTTGC-TTTTCTGCT 1605
Db 1482 AAGAG-ATGATGTCCCTTTTCTAGGATGGGTAGTGGTGGGAGTCTGCTTTTCTGCT 1540
QY 1606 GGTGGCTCTTTTGAAGTAAAAACCTGCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1665
Db 1541 GTTGGCGCGCTTGAANAATGGAATCATGTTAGTATGAGTTCCCTTGTCTCATTTGTAANA 1600
```



Query Match		48.3%;	Score 1092.8;	DB 9;	Length 1669;
Best Local Similarity		75.1%;	Pred. No. 4e-288;		
Matches 1662;		Conservative	0;	Mismatches	7;
				Indels	545;
				Gaps	2;
QY	47	GTATTGAGCGGGAGTCTAGTTCTCTGAGCGAGTGGACCCGGCAGCGGGCGATAGGGGGGCCA	106		
DB	1	GTATTGAGCGGGAGTCTAGTTCTCTGAGCGAGTGGACCCGGCAGCGGGCGATAGGGGGGCCA	60		
QY	107	GGTGCCTCCACAGTCAAGCCATGCGCTACGCGGCTCGCTACGCGGGCTGGACGACGACAG	166		
DB	61	GGTGCCTCCACAGTCAAGCCATGCGCTACGCGGCTCGCTACGCGGGCTGGACGACGACAG	120		
QY	167	TGAGGACGAGTGCCTCCGGGCTGGAGGAGAGAACCCACCAAGGACGGCTGGGTTTACTA	226		
DB	121	TGAGGACGAGTGCCTCCGGGCTGGAGGAGAGAACCCACCAAGGACGGCTGGGTTTACTA	180		
QY	227	CGCCATATCACCGAGGAGAGAGTCTAGTGGGACATCCAAATCTGGAAAGAAAGAACG	286		
DB	181	CGCCATATCACCGAGGAGAGAGTCTAGTGGGACATCCAAATCTGGAAAGAAAGAACG	240		
QY	287	AGTGGCAGGAGATTTCGCCATACGATGGGAAACAAAGAACTGATGAAACGGACAAAGTGT	346		
DB	241	AGTGGCAGGAGATTTCGCCATACGATGGGAAACAAAGAACTGATGAAACGGACAAAGTGT	300		
QY	347	TTTTGTGACCATATAAATAAAGBAACCACTTCTGAGCCCAAGACTGGCGTTTACTGT	406		
DB	301	TTTTGTGACCATATAAATAAAGBAACCACTTCTGAGCCCAAGACTGGCGTTTACTGT	360		
QY	407	GGATGATATCCGACCAAGCCAAACACCCGGCAAGATACGACGGCAGACCACTGCCAT	466		
DB	361	GGATGATATCCGACCAAGCCAAACACCCGGCAAGATACGACGGCAGACCACTGCCAT	420		
QY	467	GGAAATTTCTCAGGGCGGGATTTCACTGCGCAAGTGGTGTGTCACCTGGAGCTAAATTC	526		
DB	421	GGAAATTTCTCAGGGCGGGATTTCACTGCGCAAGTGGTGTGTCACCTGGAGCTAAATTC	480		
QY	527	AGGAATAGGGTTTGGAAACCCCAAGTCTTTTGGCCCTCCATGGTGGACATGTGATCTTGGC	586		
DB	481	AGGAATAGGGTTTGGAAACCCCAAGTCTTTTGGCCCTCCATGGTGGACATGTGATCTTGGC	540		
QY	587	CTGCAGGAATATGGCAAGCGGAGTGAAGCAGTGTACCCATTTTGAAGAATGGCATAA	646		
DB	541	CTGCAGGAATATGGCAAGCGGAGTGAAGCAGTGTACCCATTTTGAAGAATGGCATAA	594		
QY	647	AGCCAAGGTAGAAGCAATGACCTCGTCCGCTAGCGTGGCTGACGATTTTGC	706		
DB	595	-----	594		
QY	707	TGAAGCATTTCAAGGCCAAGATGTGCCTCTTTCATGTGTTGTGTGCAACGCGACAACTTT	766		
DB	595	-----	594		
QY	767	TGCTCTACCTTGGAGTCTCACCAAGATGGCTTGGAGACCACTTTCAAGTGAATCATCT	826		
DB	595	-----	594		
QY	827	GGGGCACTTCTACCTTGTCCAGTCTCTCCAGGATGTTTGTGGCGCTCAGCTCCTGCCCG	886		
DB	595	-----	594		
QY	887	TGTCATTTGGTCTCCTCAGATGCCCATCGATTTACAGATATTAACGACTCCTTGGGAA	946		
DB	595	-----	594		
QY	947	ACTGGACTTCAGTGGCTCTCTCCAAACAAAACGACTATTGGCGCATGCTGGCTTATTA	1006		
DB	595	-----	594		
QY	1007	CAGGTCCAAGCTCTGCAACATCCTCTTCTCCAAGAGTGCACCGTCGCTCTCCCCACG	1066		
DB	595	-----	594		
QY	1067	CGGGGTACGTGCGAAGCGAGTGCATCTCTGGAAATATGATGTACTCCAACATTTATCGCAG	1126		
DB	595	-----	594		
QY	1127	CTGGTGGGTGTACACACTGCTGTTTACCTTGGGAGGCCCTTTTACCAAGTCCATTCGAACA	1186		
DB	595	-----	600		
QY	1187	GGGAGCTGCCACACCGTGTACTGTCTCTCTGCCAGAACTGGAGGGTCTTGGAGGGAT	1246		
DB	601	GGGAGCTGCCACACCGTGTACTGTCTCTCTGCCAGAACTGGAGGGTCTTGGAGGGAT	660		
QY	1247	GTACTTCAACAACTGCTGCGCTGCAATGCCCTTCAACAGAGCTCAGAGCGAAGAGACGGC	1306		
DB	661	GTACTTCAACAACTGCTGCGCTGCAATGCCCTTCAACAGAGCTCAGAGCGAAGAGACGGC	720		
QY	1307	CGGGACCTGTGGGCGCTCAGCGAGAGGTGATCAAGAACCGCTTGGCAGCGACTCCCG	1366		
DB	721	CGGGACCTGTGTGGGCGCTCAGCGAGAGGTGATCAAGAACCGCTTGGCAGCGACTCCCG	780		
QY	1367	CTAAGTGGAGCTCAGAGCGGATGGGCACACACCGCCCTGTGTGTGTCCTTCAACGCA	1426		
DB	781	CTAAGTGGAGCTCAGAGCGGATGGGCACACACCGCCCTGTGTGTGTCCTTCAACGCA	840		
QY	1427	AGTGCACAGGGTGGGCGCTTCCAAATGTCTTCAACACAGATCCGCAAGAGTAAAGGA	1486		
DB	841	AGTGCACAGGGTGGGCGCTTCCAAATGTCTTCAACACAGATCCGCAAGAGTAAAGGA	900		
QY	1487	ATAAGAGCAGTCAACAAGAGTGAATAATCTTAAGTACCAATGGGAGCAGGGAATTC	1546		
DB	901	ATAAGAGCAGTCAACAAGAGTGAATAATCTTAAGTACCAATGGGAGCAGGGAATTC	960		
QY	1547	TGGGGTAAAGTATCACTTTTCTGGGCTGGGTAGGCATAGGTCTCTTTGCTTTCTGGTG	1606		
DB	961	TGGGGTAAAGTATCACTTTTCTGGGCTGGGTAGGCATAGGTCTCTTTGCTTTCTGGTG	1015		
QY	1607	GTGGCTGTGTGAAAGTAAACCTGTGTGTGTGTAGTTCCTGTATCTTCTCGTGGAGAG	1666		
DB	1016	GTGGCTGTGTGAAAGTAAACCTGTGTGTGTGTAGTTCCTGTATCTTCTCGTGGAGAG	1075		
QY	1667	CACAGCAATTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCT	1726		
DB	1076	CACCAAGATTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCT	1135		
QY	1727	AGCTACCAACCGGCGCACCTGTGACCGGGGGCTGGCCCTCTCTTCTTCTTCTTCTTCT	1786		
DB	1136	AGCTACCAACCGGCGCACCTGTGACCGGGGGCTGGCCCTCTCTTCTTCTTCTTCTTCT	1195		
QY	1787	AAGCAAGTGTTCATGTCTCTTGTGCAATGATCCAGGAGATTAATTTTCTTCTTCTTCT	1846		
DB	1196	AAGCAAGTGTTCATGTCTCTTGTGCAATGATCCAGGAGATTAATTTTCTTCTTCTTCT	1255		
QY	1847	GACCAAGCTGAGCAGCTTAGCACTGTGGGAGACAAATCTCAGACCTTGTGCCAG	1906		
DB	1256	GACCAAGCTGAGCAGCTTAGCACTGTGGGAGACAAATCTCAGACCTTGTGCCAG	1315		
QY	1907	CCAGTGAGGATGACAGTGCACCCAGAGGAGTAGAATAACGAGAACTACCAAGTGGCAA	1966		
DB	1316	CCAGTGAGGATGACAGTGCACCCAGAGGAGTAGAATAACGAGAACTACCAAGTGGCAA	1375		
QY	1967	AGTACTTGTATPAGACTCCTTTGTCTTAATGCTATGCAAAAAATTTCTTGAAGATTAACA	2026		
DB	1376	AGTACTTGTATPAGACTCCTTTGTCTTAATGCTATGCAAAAAATTTCTTGAAGATTAACA	1435		
QY	2027	AATTTTTCATATCTTCTTAGATACCTTGAAGCAGGAGGAGGAGGATATATCTTAA	2086		
DB	1436	AATTTTTCATATCTTCTTAGATACCTTGAAGCAGGAGGAGGAGGATATATCTTAA	1495		
QY	2087	GAATACACAGGATATTTTGGGGGCGAGAGATTAACCGTGTAGTTAATCCCTTGTCTGTC	2146		
DB	1496	GAATACACAGGATATTTTGGGGGCGAGAGATTAACCGTGTAGTTAATCCCTTGTCTGTC	1555		
QY	2147	AATCAGTCTCAGTCTCTTGTCTTTCATTTGTAATGTAATTAACCTCTCTGTGTGCTCGCA	2206		





Db 1650 CAGCTGCCACCACTGCCAATTTGCCACCCAAAGTTGGCTTTCTCTGTTAATGGAAGAA 1709  
Qy 1786 AAAGCA-----AGTGTTCACCTGCTCTGCTGATTCATTCAGGAGATTAATGTT 1835  
Db 1710 AAAAAAACACACACATTTCTCTTCTTCCAGCTGAGTTCAATCTCTGGAGCCATGTT 1769  
Qy 1836 TCATTCATCTCGACCAAGATGACCCAGCTTTCAGCACTGCTGGGGAGACAAATCTCGAA 1895  
Db 1770 TCACATTCCTTACCAAGTCTGACGAGCGTGGCACTTTCAGAG-----AAAACTCAGTA 1825  
Qy 1896 CTTGTGCTCC-----AGCCAGTGGAGTATGACAGTGCACCCAGGAGGAGTAGAATA 1945  
Db 1826 CTTAGTCCCAAGAGAACCAATGCGACCGTGCCATCAGAGAGGAGCAGTCTGCTGGTG 1885  
Qy 1946 CGCAGAACTACCAAGTGGC--AAAGTACTTGTCTATAGACTCTCTTGTCTAATGCTATGCAA 2003  
Db 1886 TGCAGAGCTGGTGTTCAGAGACACATTCCTGACCTTCCTTGCCTGCAAAAGTTAAACAA 1945  
Qy 2004 AAAATCTTTAGAGATTATAACAAAT---TTTCAAAATCATTCCTTAGATACCTTGAAG 2060  
Db 1946 GCAATCTTTAGAGATTATAACGAGTGCATTTTACATTTTGCAAAACCATTCCTTAGGCAG 2005  
Qy 2061 GCAGGAGGAGGAGTATATACCTTAAAGATACACAGGATATTTGGGGGCGCAGAGATAA 2120  
Db 2006 GCAAGGAGGAGGAGCAGCTGCTCTTAAAGATACATGGGATATGCTGGAGGTGGGAAGCAAG 2065  
Qy 2121 AAGCTTAGTTAAATCCCTTTGCTGTCTCA 2147  
Db 2066 GAAGTAGTTCCTCCAGCTCTTCACTCA 2092

RESULT 14  
AC092376 168083 bp DNA linear PRI 15-MAR-2003  
LOCUS Homo sapiens chromosome 16 clone RP11-679B19, complete sequence.  
DEFINITION AC092376 AC036142  
ACCESSION AC092376.3 GI:28973808  
VERSION HTG.  
KEYWORDS Homo sapiens (human)  
SOURCE ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 168083)  
AUTHORS DOE Joint Genome Institute, Stanford Human Genome Center and Los  
Alamos National Laboratory.  
TITLE Direct Submission  
JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 168083)  
AUTHORS DOE Joint Genome Institute.  
JOURNAL Direct Submission  
REFERENCE 3 (bases 1 to 168083)  
AUTHORS DOE Joint Genome Institute.  
JOURNAL Submitted (03-JUL-2001) Production Sequencing Facility, DOE Joint  
Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA  
REFERENCE 4 (bases 1 to 168083)  
AUTHORS DOE Joint Genome Institute, Stanford Human Genome Center and Los  
Alamos National Laboratory.  
JOURNAL Direct Submission  
REFERENCE 5 (bases 1 to 168083)  
AUTHORS DOE Joint Genome Institute, Stanford Human Genome Center and Los  
Alamos National Laboratory.  
JOURNAL Submitted (15-MAR-2003) DOE Joint Genome Institute, 2800 Mitchell  
Drive, Walnut Creek, CA 94598, USA  
COMMENT On Mar 15, 2003 this sequence version replaced gi:15187320.  
Draft Sequence Produced by DOE Joint Genome Institute  
www.jgi.doe.gov  
Finishing Completed at Stanford Human Genome Center and Los Alamos  
National Laboratory  
www.shgc.stanford.edu  
Quality: Phrap Quality >=40 100% of Sequence;  
Estimated Total Number of Errors is 0.  
FEATURES  
Location/Qualifiers  
1..168083

/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/chromosome="16"  
/clone="RP11-679B19"  
ORIGIN  
Query Match 46.9% Score 1062.2; DB 9; Length 168083;  
Best Local Similarity 99.3%; Pred. No. 2.8e-279;  
Matches 1067; Conservative 0; Mismatches 8; Indels 0; Gaps 0;  
Qy 1167 TTCAACAAGTCCATGCAACAGGAGCTGCCACCAACCGCTGTAATCTGCTGCTGCCAGAA 1226  
Db 23701 TTCTTGGATTTCAGCAACAGGAGCTGCCACCAACCGTGTACTGTGTGCTGCTGCCAGAA 23760  
Qy 1227 CTGAGAGGCTCTGGAGGAGTGTACTTCAACAACTGTGCGCTGCATGCCCTCAACGAA 1286  
Db 23761 CTGAGAGGCTCTGGAGGAGTGTACTTCAACAACTGTGCGCTGCATGCCCTCAACGAA 23820  
Qy 1287 GCTCAGAGCAAGAGAGCGCGGACCTGTGGCGCTCAGCGAGAGGCTGATCCAGAA 1346  
Db 23821 GCTCAGAGCAAGAGAGCGCGGACCTGTGGCGCTCAGCGAGAGGCTGATCCAGAA 23880  
Qy 1347 CGCTTTGGCAGCAGTCCGGCTAAGTGGAGCTCAGAGCGGATGGGCAACACACACCGGCC 1406  
Db 23881 CGCTTTGGCAGCAGTCCGGCTAAGTGGAGCTCAGAGCGGATGGGCAACACACACCGGCC 23940  
Qy 1407 TGTGTGTCTCCCTCAGCAAGTGCAGAGGCTGGGCGCTTCCAAATGTCCCTCCACAC 1466  
Db 23941 TGTGTGTCTCCCTCAGCAAGTGCAGAGGCTGGGCGCTTCCAAATGTCCCTCCACAC 24000  
Qy 1467 AGATCCGCAAGAGTAAAGNAATAGAGAGCTCACAACAGAGTGAATAATCTTAAGTACC 1526  
Db 24001 AGATCCGCAAGAGTAAAGNAATAGAGAGCTCACAACAGAGTGAATAATCTTAAGTACC 24060  
Qy 1527 AATGGAAGCAGGGAATTCCTGGGGTAAAGTATCATCTTTCTGGGGCTGGGCTAGGCATA 1586  
Db 24061 AATGGAAGCAGGGAATTCCTGGGGTAAAGTATCATCTTTCTGGGGCTGGGCTAGGCATA 24120  
Qy 1587 GGTCTCTTTCTTTCTGGTGGCTGTGTTGAAGTAAAAACCTGCTGGTGTAGGT 1646  
Db 24121 GGTCTCTTTCTTTCTGGTGGCTGTGTTGAAGTAAAAACCTGCTGGTGTAGGT 24180  
Qy 1647 TCGGTATCTCCCTGGAGAGCACCAGCAATTCCTTTCTTTTACTGTTATAGAAATAGCCT 1706  
Db 24181 TCGGTATCTCCCTGGAGAGCACCAGCAATTCCTTTCTTTTACTGTTATAGAAATAGCCT 24240  
Qy 1707 GAGGTCCCTCTGCTCCATCCAGTACCAACAGGCGCACCACTGAGCGGGGCTGGCCT 1766  
Db 24241 GAGGTCCCTCTGCTCCATCCAGTACCAACAGGCGCACCACTGAGCGGGGCTGGCCT 24300  
Qy 1767 TCTCTTACTTAGGAAGAAAAAGAGTGTTCATCTGCTCTCTGCTGCATTTGATCCAGGAG 1826  
Db 24301 TCTCTTACTTAGGAAGAAAAAGAGTGTTCATCTGCTCTCTGCTGCATTTGATCCAGGAG 24360  
Qy 1827 ATATATGTTTCAATTCATCTGACCAAGACTGAGCCAGCTTAGCAATGCTGGGGACAA 1886  
Db 24361 ATATATGTTTCAATTCATCTGACCAAGACTGAGCCAGCTTAGCAATGCTGGGGACAA 24420  
Qy 1887 ATCTCAGAACTTGTCTCCAGCCAGTGAAGTACAGTGCACCCAGAGGAGTAGAATAC 1946  
Db 24421 ATCTCAGAACTTGTCTCCAGCCAGTGAAGTACAGTGCACCCAGAGGAGTAGAATAC 24480  
Qy 1947 GCAGAACTACCAAGTGGCAAGTACTGTTCATAGACTCTCTTTGCTTAATCTATGCAAAA 2006  
Db 24481 GCAGAACTACCAAGTGGCAAGTACTGTTCATAGACTCTCTTTGCTTAATCTATGCAAAA 24540  
Qy 2007 ATTCTTTAGAGATTATAACAAATTTTCAAAATCTTCCCTTAGATACCTTGAAGGAGGA 2066  
Db 24541 ATTCTTTAGAGATTATAACAAATTTTCAAAATCTTCCCTTAGATACCTTGAAGGAGGA 24600  
Qy 2067 AGGGAAGCGTATATATCTTAAAGAAATACAGAGGATATTTTGGGGGCGCAGAAATAACGTT 2126









GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 24, 2004, 09:49:46 ; Search time 58 Seconds  
(without alignments)  
2016.805 Million cell updates/sec

Title: US-09-978-318B-1

Perfect score: 2203

Sequence: 1 MAALRYAGLDDTSEDELPP.....RTLWALSERLIQERIGSQSG 414

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A\_Geneseq\_29Jan04.\*

1: Geneseq1980s.\*

2: Geneseq1990s.\*

3: Geneseq2000s.\*

4: Geneseq2001s.\*

5: Geneseq2002s.\*

6: Geneseq2003as.\*

7: Geneseq2003bs.\*

8: Geneseq2004s.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2203	100.0	414	5	ABG96559 Human sho
2	2196	99.7	414	4	AAB93119 Human pro
3	2196	99.7	414	4	AAB85886 Human gen
4	2191	99.5	414	4	AAB84485 Amino aci
5	2190	99.4	415	4	AAB18308 Human end
6	2184	99.1	423	4	ABB10203 Human CDN
7	2184	99.1	423	5	ABP66790 Human pol
8	1869	84.8	363	3	AAB84484 Amino aci
9	1023	46.4	409	4	ABB63100 Drosophil
10	927	42.1	189	4	AAB84486 Amino aci
11	862.5	39.2	173	4	ABB10434 Human CDN
12	862.5	39.2	173	4	AAB18436 Human end
13	862.5	39.2	173	5	ABP67021 Human pol
14	550.5	25.0	316	5	AAE14438 Human dru
15	549.5	24.9	316	6	ADA54192 Human pro
16	549.5	24.9	316	7	ABR82985 Human unn
17	540	24.5	321	5	ABP69268 Human pol
18	533.5	24.2	322	3	AAG48405 Arabidops
19	527	23.9	323	2	AAW15759 Cotton fi
20	520.5	23.6	320	3	AAG05338 Arabidops
21	519.5	23.6	296	4	ABB58831 Drosophil
22	514.5	23.4	316	3	AAG48393 Arabidops
23	513.5	23.3	318	3	AAY82014 Human imm
24	513.5	23.3	318	3	AAB12158 Hydrophob
25	513.5	23.3	318	4	AAW01129 Human pro

## ALIGNMENTS

RESULT 1

ABG96559 standard; protein; 414 AA.

XX AC ABG96559;

XX DT 12-DEC-2002 (first entry)

XX DE Human short chain dehydrogenase family member WWOX.

XX Short chain dehydrogenase; SDR; human; antimicrotoca; peatidice;  
herbicide; DHPH deficiency; phenylketonuria; galactosaemia III;  
dienoyl CoA reductase deficiency; adrenal hyperplasia; ovarian cancer;  
adrenogenital syndrome; mineralcorticoid excess syndrome; breast cancer;  
male pseudohemaphroditism; Zellweger syndrome; Down's syndrome;  
polycystic kidney disease; Alzheimer's disease; retinitis pigmentosa;  
retinitis punctata albescens; arterial hypertension; follicular lymphoma;  
hepatocarcinogenesis; fungicide; antibiotic.

XX OS Homo sapiens.

XX PN W0200212544-A2.

XX PD 14-FEB-2002.

XX PF 07-AUG-2001; 2001WO-EP009140.

XX PR 07-AUG-2000; 2000US-0223436P.

XX PA (BION-) BIONETWORKS GMBH.

XX PI Wilckens T;

XX DR WPI; 2002-241770/29.

XX Identifying or verifying members of the short chain dehydrogenase (SDR) family, useful for novel drug development (e.g. for the development of antimicrotoca, pesticides or herbicides), by employing an algorithm using core SDR motifs.

XX Disclosure; Fig 4; 168pp; English.

XX The invention relates to identifying or verifying members of the short chain dehydrogenase (SDR) family comprising employing an algorithm using core SDR motifs (MFI-MR4 and MVL, MVL given in the specification) for searching members of the SDR family. Also included are a member of the SDR family identified with the method above, a method for providing modulators for members of the SDR family, a method for evaluation of lead

AAM93777 Human pol  
AAU69774 Human pro  
AAB74812 Prostate  
AAG99014 Human pro  
ABU71665 Prostate  
ABG96558 Human sho  
ABB95234 Human P50  
ABR54346 Prostate  
ADB75403 Prostate  
ADB13789 Human pro  
AAY48616 Human bre  
AAB36900 Human ARS  
ABR63833 Human sho  
ABR82986 Human CGI  
ABB58826 Drosophil  
ASG81281 Human AFP  
ABB66596 Drosophil  
AAY41730 Human PRO  
AAB44286 Human PRO  
AAU29114 Human PRO

26 513.5 23.3 318 4 AAM93777  
27 513.5 23.3 318 4 AAU69774  
28 513.5 23.3 318 4 AAB74812  
29 513.5 23.3 318 4 AAG99014  
30 513.5 23.3 318 4 ABU71665  
31 513.5 23.3 318 5 ABG96558  
32 513.5 23.3 318 5 ABB95234  
33 513.5 23.3 318 6 ABR54346  
34 513.5 23.3 318 7 ADB75403  
35 513.5 23.3 318 7 ADB13789  
36 513.5 23.3 333 2 AAY48616  
37 512.5 23.3 314 3 AAB36900  
38 512.5 23.3 318 7 ABR63833  
39 512.5 23.3 318 7 ABR82986  
40 510.5 23.2 300 4 ABB58826  
41 508 23.1 296 4 AAG81281  
42 507 23.0 594 4 ABB66596  
43 505 22.9 336 2 AAY41730  
44 505 22.9 336 3 AAB44286  
45 505 22.9 336 4 AAU29114

CC -candidates for possible modulators of a member of the SDR family and a  
 CC method for detecting clinically relevant polymorphisms or single  
 CC nucleotide polymorphisms. The method is useful for screening SDR  
 CC sequences and modulators of the SDR family. The method is especially  
 CC useful as a platform for novel drug development. The SDRs can serve for  
 CC the development of e.g. antimicrobials, pesticides or herbicides. The  
 CC modulators may be especially useful for the prophylaxis, treatment or/and  
 CC diagnosis of diseases (e.g. DHRP deficiency, phenylketonuria, diacyl CoA  
 CC reductase deficiency, galactosaemia III, adrenal hyperplasia,  
 CC adrenogenital syndrome, mineralocorticoid excess syndrome, ovarian cancer,  
 CC breast cancer, male pseudohypoparathyroidism, Zellweger syndrome, polycystic  
 CC kidney disease, Alzheimer's disease, retinitis punctata albescens,  
 CC retinitis pigmentosa, Down's syndrome, arterial hypertension, follicular  
 CC lymphoma and hepatocarcinogenesis) particularly as a fungicide or  
 CC antibiotic. The present sequence is one of 39 human SDR family members  
 CC identified by the method of the invention  
 XX  
 XX

XX SQ Sequence 414 AA;

Query Match 100.0%; Score 2203; DB 5; Length 414;  
 Best Local Similarity 100.0%; Pred. No. 1e-215;  
 Matches 414; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Qy 1 MAALRYAGLDDTDSDELPPGWEERTTKDGWVYVYANHTTEKTQWEHPKTKGRKRVAGDLP 60  
 Db 1 MAALRYAGLDDTDSDELPPGWEERTTKDGWVYVYANHTTEKTQWEHPKTKGRKRVAGDLP 60  
 Qy 61 YGWEQETDENGQVFVDHINKRTTYLDPRLAFTVDDNPTKTTTQRYDGTSTAMEILQGR 120  
 Db 61 YGWEQETDENGQVFVDHINKRTTYLDPRLAFTVDDNPTKTTTQRYDGTSTAMEILQGR 120  
 Qy 121 DFTGKVVVVTGANGSIGFETAKSPALHGAHVILACRNMARASEAVSRILEBWHKAKVEAM 180  
 Db 121 DFTGKVVVVTGANGSIGFETAKSPALHGAHVILACRNMARASEAVSRILEBWHKAKVEAM 180  
 Qy 181 TLDLALLRSVOHFAPAEAFKAKNVPLHLVLCNAATFALPWSLTKDGLTTFQVNHGFLYV 240  
 Db 181 TLDLALLRSVOHFAPAEAFKAKNVPLHLVLCNAATFALPWSLTKDGLTTFQVNHGFLYV 240  
 Qy 241 QLLQDVLCRSAPARVIVVSSSHRFTDINDSLGKLDPSRLSPTKNDYWMALYNRSLKCN 300  
 Db 241 QLLQDVLCRSAPARVIVVSSSHRFTDINDSLGKLDPSRLSPTKNDYWMALYNRSLKCN 300  
 Qy 301 ILFNSNELHRRILSPRGVTSNAVHPGNMYSNTHRSWWVYTLFTLARPTKSMQGAATTV 360  
 Db 301 ILFNSNELHRRILSPRGVTSNAVHPGNMYSNTHRSWWVYTLFTLARPTKSMQGAATTV 360  
 Qy 361 YCAAVPELEGIGMYFNNCCRCMPSPQAQSEETARTLWALSERLIQERLGSQS 414  
 Db 361 YCAAVPELEGIGMYFNNCCRCMPSPQAQSEETARTLWALSERLIQERLGSQS 414

RESULT 2  
 AAB93119  
 ID AAB93119 standard; protein; 414 AA.

XX AC AAB93119;

XX DT 26-JUN-2001 (first entry)

XX DE Human protein sequence SEQ ID NO:11995.

XX KW Human; primer; detection; diagnosis; antisense therapy; gene therapy.

XX OS Homo sapiens.

XX PN EP1074617-A2.

XX PD 07-FEB-2001.

XX PF 28-JUL-2000; 2000EP-00116126.

XX PR 29-JUL-1999; 99JP-00248036.

PR 27-AUG-1999; 99JP-00300253.  
 PR 11-JAN-2000; 2000JP-00118776.  
 PR 02-MAY-2000; 2000JP-00183767.  
 PR 09-JUN-2000; 2000JP-00241899.  
 XX (HELI-) HELIX RES INST.  
 XX Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;  
 XX Iehi S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;  
 XX WPI; 2001-318749/34.  
 XX Primer sets for synthesizing polynucleotides, particularly the 5602 full-  
 XX length cDNAs defined in the specification, and for the detection and/or  
 XX diagnosis of the abnormality of the proteins encoded by the full-length  
 XX cDNAs.

Claim 8; SEQ ID NO 11995; 2537pp + Sequence Listing; English.

The present invention describes primer sets for synthesizing 5602 full-  
 length cDNAs defined in the specification. Where a primer set comprises:  
 (a) an oligo-dr primer and an oligonucleotide complementary to the  
 complementary strand of a polynucleotide which comprises one of the 5602  
 nucleotide sequences defined in the specification, where the  
 oligonucleotide comprises at least 15 nucleotides; or (b) a combination  
 of an oligonucleotide comprising a sequence complementary to the  
 complementary strand of a polynucleotide which comprises a 5'-end  
 sequence and an oligonucleotide comprising a sequence complementary to a  
 polynucleotide which comprises a 3'-end sequence, where the  
 oligonucleotide comprises at least 15 nucleotides and the combination of  
 the 5'-end sequence/3'-end sequence is selected from those defined in the  
 specification. The primer sets can be used in antisense therapy and in  
 gene therapy. The primers are useful for synthesizing polynucleotides,  
 particularly full-length cDNAs. The primers are also useful for the  
 detection and/or diagnosis of the abnormality of the proteins encoded by  
 the full-length cDNAs. The primers allow obtaining of the full-length  
 cDNAs easily without any specialised methods. AAB93116 to AAB93628 and  
 AAB93629 to AAB93742 represent human cDNA sequences; AAB92446 to AAB95893  
 represent human amino acid sequences; and AAB93629 to AAB93632 represent  
 oligonucleotides, all of which are used in the exemplification of the  
 present invention

Sequence 414 AA;

Query Match 99.7%; Score 2196; DB 4; Length 414;  
 Best Local Similarity 99.8%; Pred. No. 5.3e-215;  
 Matches 413; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MAALRYAGLDDTDSDELPPGWEERTTKDGWVYVYANHTTEKTQWEHPKTKGRKRVAGDLP 60  
 Db 1 MAALRYAGLDDTDSDELPPGWEERTTKDGWVYVYANHTTEKTQWEHPKTKGRKRVAGDLP 60  
 Qy 61 YGWEQETDENGQVFVDHINKRTTYLDPRLAFTVDDNPTKTTTQRYDGTSTAMEILQGR 120  
 Db 61 YGWEQETDENGQVFVDHINKRTTYLDPRLAFTVDDNPTKTTTQRYDGTSTAMEILQGR 120  
 Qy 121 DFTGKVVVVTGANGSIGFETAKSPALHGAHVILACRNMARASEAVSRILEBWHKAKVEAM 180  
 Db 121 DFTGKVVVVTGANGSIGFETAKSPALHGAHVILACRNMARASEAVSRILEBWHKAKVEAM 180  
 Qy 181 TLDLALLRSVOHFAPAEAFKAKNVPLHLVLCNAATFALPWSLTKDGLTTFQVNHGFLYV 240  
 Db 181 TLDLALLRSVOHFAPAEAFKAKNVPLHLVLCNAATFALPWSLTKDGLTTFQVNHGFLYV 240  
 Qy 241 QLLQDVLCRSAPARVIVVSSSHRFTDINDSLGKLDPSRLSPTKNDYWMALYNRSLKCN 300  
 Db 241 QLLQDVLCRSAPARVIVVSSSHRFTDINDSLGKLDPSRLSPTKNDYWMALYNRSLKCN 300  
 Qy 301 ILFNSNELHRRILSPRGVTSNAVHPGNMYSNTHRSWWVYTLFTLARPTKSMQGAATTV 360  
 Db 301 ILFNSNELHRRILSPRGVTSNAVHPGNMYSNTHRSWWVYTLFTLARPTKSMQGAATTV 360  
 Qy 361 YCAAVPELEGIGMYFNNCCRCMPSPQAQSEETARTLWALSERLIQERLGSQS 414



CC associated with the tumour, a breakpoint of one or more chromosomal  
CC rearrangements associated with the tumour and a pause site within the  
CC FRA16 gene. FOR nucleic acid molecules are useful as markers to identify  
CC relationship between the fragile site (FRA16D) and the DNA instability in  
CC neoplasia which allows better diagnosis of cancers associated with the  
CC region  
XX  
SQ Sequence 414 AA;

Query Match 99.5%; Score 2191; DB 4; Length 414;  
Best Local Similarity 99.5%; Pred. No. 1.7e-214;  
Matches 412; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 MAALRYAGLDDTDSDELPPGWEERTTKDGVVYANHTTEKTQWEHPKTKGRKRVAGDLP 60  
DB 1 MAALRYAGLDDTDSDELPPGWEERTTKDGVVYANHTTEKTQWEHPKTKGRKRVAGDLP 60

QY 61 YGWEQETDENGQVFFVDHINKRTTYLDPLAFTVDNDTKPTTRQRYDGTSTAMEILQGR 120  
DB 61 YGWEQETDENGQVFFVDHINKRTTYLDPLAFTVDNDTKPTTRQRYDGTSTAMEILQGR 120

QY 121 DFTGKVVVVTGANSIGIGETAKSPALHGAHVILACRNWARASEAVSRILEEWHKAKVEAM 180  
DB 121 DFTGKVVVVTGANSIGIGETAKSPALHGAHVILACRNWARASEAVSRILEEWHKAKVEAM 180

QY 181 TIDLALLRSVQHFAFAKAKVPLHLVCNATPALPWSLTKDGLTTFQVNHHLGHFYL 240  
DB 181 TIDLALLRSVQHFAFAKAKVPLHLVCNATPALPWSLTKDGLTTFQVNHHLGHFYL 240

QY 241 QLLQDVLCSAPARIVIVSSSHRFTDINDSLGKLDLFRSLTKNDYAMLAYNRSKLCN 300  
DB 241 QLLQDVLCSAPARIVIVSSSHRFTDINDSLGKLDLFRSLTKNDYAMLAYNRSKLCN 300

QY 301 ILFSNELHRRISPRGVTSNAVHPGNMYSNHRSMWVYTLFTLARPPTKSMQOQAATTV 360  
DB 301 ILFSNELHRRISPRGVTSNAVHPGNMYSNHRSMWVYTLFTLARPPTKSMQOQAATTV 360

QY 361 YCAAPELEGLGGMVFNCCRCMPSPQAQSEBTARTLWALSERLIQERLGSQSG 414  
DB 361 YCAAPELEGLGGMVFNCCRCMPSPQAQSEBTARTLWALSERLIQERLGSQSG 414

RESULT 5  
AAU18308  
ID AAU18308 standard; protein; 415 AA.  
AC AAU18308;  
XX  
DT 21-NOV-2001 (first entry)  
XX  
DE Human endocrine polypeptide SEQ ID No 263.  
XX  
KW Endocrine protein; human; mouse; rabbit; goat; horse; food additive; cat;  
KW dog; chicken; sheep; immunosuppressive; antiarthritic; vasotropic;  
KW antihemetic; antiproliferative; cytostatic; cardiac; neuroprotective;  
KW cerebroprotective; nootropic; antibacterial; virucide; fungicide; cancer;  
KW ophthalmological; vulnary; gene therapy; autoimmune disease; neoplasm;  
KW hyperproliferative disorder; breast; liver; cardiovascular disorder;  
KW cerebrovascular disorder; nervous system disorder; bacterial infection;  
KW fungal infection; viral infection; ocular disorder; endocrine disorder;  
KW gastrointestinal disorder; renal disorder; respiratory disorder;  
KW wound healing; skin aging; organ transplantation; food preservative;  
KW tissue regeneration; anti-infertility.  
XX  
OS Homo sapiens.  
XX  
XX WO20015364-A2.  
PN  
PD  
XX  
XX 02-AUG-2001.  
XX  
XX 17-JAN-2001; 2001WO-US001308.  
PF  
XX  
XX 31-JAN-2000; 2000US-0179065P.  
PR





XX	Homo sapiens.	26-SEP-2000;	2000US-0235484P	PR
XX	XX	27-SEP-2000;	2000US-0235834P	PR
PN	WO200154474-A2.	27-SEP-2000;	2000US-0235836P	PR
XX		29-SEP-2000;	2000US-0236327P	PR
XX		29-SEP-2000;	2000US-0236367P	PR
PD		29-SEP-2000;	2000US-0236369P	PR
XX	02-AUG-2001.	29-SEP-2000;	2000US-0236369P	PR
XX		29-SEP-2000;	2000US-0236370P	PR
XX		02-OCT-2000;	2000US-0236802P	PR
XX	17-JAN-2001; 2001WO-US0001349.	02-OCT-2000;	2000US-0237037P	PR
XX		02-OCT-2000;	2000US-0237038P	PR
PR	31-JAN-2000;	02-OCT-2000;	2000US-0237039P	PR
PR	04-FEB-2000;	02-OCT-2000;	2000US-0237040P	PR
PR	24-FEB-2000;	02-OCT-2000;	2000US-0239335P	PR
PR	02-MAR-2000;	13-OCT-2000;	2000US-0239337P	PR
PR	16-MAR-2000;	13-OCT-2000;	2000US-0240960P	PR
PR	17-MAR-2000;	20-OCT-2000;	2000US-0240960P	PR
PR	18-APR-2000;	20-OCT-2000;	2000US-0241221P	PR
PR	19-MAY-2000;	20-OCT-2000;	2000US-0241785P	PR
PR	07-JUN-2000;	20-OCT-2000;	2000US-0241786P	PR
PR	28-JUN-2000;	20-OCT-2000;	2000US-0241787P	PR
PR	30-JUN-2000;	20-OCT-2000;	2000US-0241808P	PR
PR	07-JUL-2000;	20-OCT-2000;	2000US-0241809P	PR
PR	07-JUL-2000;	20-OCT-2000;	2000US-0241826P	PR
PR	11-JUL-2000;	01-NOV-2000;	2000US-0244617P	PR
PR	11-JUL-2000;	08-NOV-2000;	2000US-0246474P	PR
PR	14-JUL-2000;	08-NOV-2000;	2000US-0246475P	PR
PR	26-JUL-2000;	08-NOV-2000;	2000US-0246476P	PR
PR	26-JUL-2000;	08-NOV-2000;	2000US-0246477P	PR
PR	14-AUG-2000;	08-NOV-2000;	2000US-0246478P	PR
PR	14-AUG-2000;	08-NOV-2000;	2000US-0246523P	PR
PR	14-AUG-2000;	08-NOV-2000;	2000US-0246524P	PR
PR	14-AUG-2000;	08-NOV-2000;	2000US-0246525P	PR
PR	14-AUG-2000;	08-NOV-2000;	2000US-0246526P	PR
PR	14-AUG-2000;	08-NOV-2000;	2000US-0246527P	PR
PR	14-AUG-2000;	08-NOV-2000;	2000US-0246528P	PR
PR	14-AUG-2000;	08-NOV-2000;	2000US-0246532P	PR
PR	14-AUG-2000;	08-NOV-2000;	2000US-0246609P	PR
PR	14-AUG-2000;	08-NOV-2000;	2000US-0246610P	PR
PR	14-AUG-2000;	08-NOV-2000;	2000US-0246611P	PR
PR	14-AUG-2000;	08-NOV-2000;	2000US-0246613P	PR
PR	18-AUG-2000;	17-NOV-2000;	2000US-0249207P	PR
PR	22-AUG-2000;	17-NOV-2000;	2000US-0249208P	PR
PR	22-AUG-2000;	17-NOV-2000;	2000US-0249209P	PR
PR	22-AUG-2000;	17-NOV-2000;	2000US-0249210P	PR
PR	22-AUG-2000;	17-NOV-2000;	2000US-0249211P	PR
PR	30-AUG-2000;	17-NOV-2000;	2000US-0249212P	PR
PR	01-SEP-2000;	17-NOV-2000;	2000US-0249213P	PR
PR	01-SEP-2000;	17-NOV-2000;	2000US-0249214P	PR
PR	01-SEP-2000;	17-NOV-2000;	2000US-0249215P	PR
PR	05-SEP-2000;	17-NOV-2000;	2000US-0249216P	PR
PR	05-SEP-2000;	17-NOV-2000;	2000US-0249217P	PR
PR	06-SEP-2000;	17-NOV-2000;	2000US-0249218P	PR
PR	06-SEP-2000;	17-NOV-2000;	2000US-0249244P	PR
PR	08-SEP-2000;	17-NOV-2000;	2000US-0249245P	PR
PR	08-SEP-2000;	17-NOV-2000;	2000US-0249264P	PR
PR	08-SEP-2000;	17-NOV-2000;	2000US-0249265P	PR
PR	08-SEP-2000;	17-NOV-2000;	2000US-0249297P	PR
PR	08-SEP-2000;	17-NOV-2000;	2000US-0249299P	PR
PR	08-SEP-2000;	17-NOV-2000;	2000US-0249300P	PR
PR	08-SEP-2000;	01-DEC-2000;	2000US-0250160P	PR
PR	12-SEP-2000;	01-DEC-2000;	2000US-0250391P	PR
PR	14-SEP-2000;	05-DEC-2000;	2000US-0251030P	PR
PR	14-SEP-2000;	05-DEC-2000;	2000US-0251988P	PR
PR	14-SEP-2000;	05-DEC-2000;	2000US-0256719P	PR
PR	14-SEP-200			



XX The invention relates to novel genes (ABV83682-ABV84101) and proteins  
 CC (ABP66710-ABP67129) useful for preventing, treating or ameliorating  
 CC medical conditions e.g. by protein or gene therapy. The genes are  
 CC isolated from a range of human tissues disclosed in the specification.  
 CC The nucleic acids, proteins, antibodies and (ant)agonists are useful in  
 CC the diagnosis, treatment and prevention of: (a) cancer, e.g. breast and  
 CC ovarian cancer and other cancers of the adrenal gland, bone, bone marrow,  
 CC breast, gastrointestinal tract, liver, lung, or urogenital; (b) immune  
 CC disorders e.g. Addison's disease, allergies, autoimmune haemolytic  
 CC anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's disease,  
 CC multiple sclerosis, rheumatoid arthritis and ulcerative colitis; (c)  
 CC cardiovascular disorders such as myocardial ischaemia; (d) wound healing  
 CC ; (e) neurological diseases e.g. cerebral anoxia and epilepsy; and (f)  
 CC infectious diseases such as viral, bacterial, fungal and parasitic  
 CC infections. Note: The sequence data for this patent did not form part of  
 CC the printed specification, but was obtained in electronic format directly  
 CC from WIPO at ftp.wipo.int/pub/published\_pct\_sequences  
 XX

XX Sequence 423 AA;

Query Match 99.1%; Score 2184; DB 5; Length 423;  
 Best Local Similarity 99.0%; Pred. No. 9.2e-214; Indels 0;  
 Matches 410; Conservative 0; Mismatches 4; Gaps 0;

QY 1 MAALRYAGLDGTTSEDELPDGEERTTKDGVVYANHTTEKTQWHPKTKGKRKRVAGDLP 60  
 DB 10 MAALRYAGLDGTTSEDELPDGEERTTKDGVVYANHTTEKTQWHPKTKGKRKRVAGDLP 69  
 QY 61 YGWEQETDENGQVFFVDHINKRTTYLDPLAFTVDNDPTKPTRQRYDGSITAMAILQGR 120  
 DB 70 YGWEQETDENGQVFFVDHINKRTTYLDPLAFTVDNDPTKPTRQRYDGSITAMAILQGR 129  
 QY 121 DFTGKVVVVTGANSIGGFETAKSFALHGAHVILACRNMARASEAVSRILEEWHKAKVETM 180  
 DB 130 DFTGKVVVVTGANSIGGFETAKSFALHGAHVILACRNMARASEAVSRILEEWHKAKVETM 189  
 QY 181 TLDLALLRSVQHFAFAKAKVPLHVLVCNATFALPWSLTKDGLTTFQVNHGHFYL 240  
 DB 190 TLDLALLRSVQHFAFAKAKVPLHVLVCNATFALPWSLTKDGLTTFQVNHGHFYL 249  
 QY 241 QLLQDVLCRSAPARVIVVSSSESHRFTDINDSLGKLDLFSRLSPTKNDYWMALAYNRSKLCN 300  
 DB 250 QLLQDVLCRSAPARVIVVSSSESHRFTDINDSLGKLDLFSRLSPTKNDYWMALAYNRSKLCN 309  
 QY 301 ILFSNELHRRILSPRGVTSNAVHPGNMYSNIHRSWVYTLFTLARPTKSMQCGAATV 360  
 DB 310 ILFSNELHRRILSPRGVTSNAVHPGNMYSNIHRSWVYTLFTLARPTKSMQCGAATV 369  
 QY 361 YCAAYPELEGLGMYFNCCRCMPSPQAQSEETARTLWALSERLIQERLGSQSG 414  
 DB 370 YCAAYPELEGLGMYFNCCRCMPSPQAQSEETARTLWALSERLIQERLGSQSG 423

RESULT 8  
 AAB84484  
 ID AAB84484 standard; protein; 363 AA.

XX AAB84484;

XX 05-SEP-2001 (first entry)

XX Amino acid sequence of FRA16D oxidoreductase (FOR) I gene cDNA.

XX Cancer associated protein; FOR gene; FRA16D; fragile site; aphidicolin;  
 KW chromosomal rearrangement; cancer; splice variant; DNA instability;  
 XX FRA16D oxidoreductase; neoplasia.

OS Homo sapiens.

XX WO200144466-A1.

XX 21-JUN-2001.

XX 15-DEC-2000; 2000WO-AU001539.  
 XX 16-DEC-1999; 99AU-00004711.  
 XX 19-APR-2000; 2000AU-00007025.  
 XX (WOME-) WOMEN'S & CHILDREN'S HOSPITAL.  
 XX Richards R, Ried K, Finnis M, Hobson L, Mangelsdorf M, Dayan S;  
 PI Nancarrow J, Woollatt E, Baker E;  
 XX WPI; 2001-398151/42.  
 XX N-PSDB; AAH27866.  
 DR Novel isolated 16q23.2 nucleic acid molecule, FRA16D oxidoreductase (FOR)  
 XX gene associated with FRA16D site, useful for early diagnosis and  
 PT assessment of risk of cancers associated with the FRA16D region.  
 XX Claim 31; Fig 9; 150pp; English.

XX The present sequence represents a human FRA16D oxidoreductase (FOR) I  
 CC transcript. The FOR gene encodes a cancer associated protein. The FRA16D  
 CC site is a fragile site induced by aphidicolin, which is located within  
 CC the FOR gene. The fragile site is the location of breakpoints of a  
 CC variety of chromosomal rearrangements, and other mutations associated  
 CC with cancers. The FOR protein is expressed as a number of splice  
 CC variants. FOR gene polynucleotide fragments are capable of acting as  
 CC specific primers or probes for detecting cancer associated variations of  
 CC DNA sequence such as a point mutation or small DNA rearrangement  
 CC associated with the tumour, a breakpoint of one or more chromosomal  
 CC rearrangements associated with the tumour and a pause site within the  
 CC FRA16 gene. FOR nucleic acid molecules are useful as markers to identify  
 CC relationship between the fragile site (FRA16D) and the DNA instability in  
 CC neoplasia which allows better diagnosis of cancers associated with the  
 CC region

XX Sequence 363 AA;

Query Match 84.8%; Score 1869; DB 4; Length 363;  
 Best Local Similarity 99.7%; Pred. No. 1.1e-181;  
 Matches 351; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MAALRYAGLDGTTSEDELPDGEERTTKDGVVYANHTTEKTQWHPKTKGKRKRVAGDLP 60  
 DB 1 MAALRYAGLDGTTSEDELPDGEERTTKDGVVYANHTTEKTQWHPKTKGKRKRVAGDLP 60  
 QY 61 YGWEQETDENGQVFFVDHINKRTTYLDPLAFTVDNDPTKPTRQRYDGSITAMAILQGR 120  
 DB 61 YGWEQETDENGQVFFVDHINKRTTYLDPLAFTVDNDPTKPTRQRYDGSITAMAILQGR 120  
 QY 121 DFTGKVVVVTGANSIGGFETAKSFALHGAHVILACRNMARASEAVSRILEEWHKAKVETM 180  
 DB 121 DFTGKVVVVTGANSIGGFETAKSFALHGAHVILACRNMARASEAVSRILEEWHKAKVETM 180  
 QY 181 TLDLALLRSVQHFAFAKAKVPLHVLVCNATFALPWSLTKDGLTTFQVNHGHFYL 240  
 DB 181 TLDLALLRSVQHFAFAKAKVPLHVLVCNATFALPWSLTKDGLTTFQVNHGHFYL 240  
 QY 241 QLLQDVLCRSAPARVIVVSSSESHRFTDINDSLGKLDLFSRLSPTKNDYWMALAYNRSKLCN 300  
 DB 241 QLLQDVLCRSAPARVIVVSSSESHRFTDINDSLGKLDLFSRLSPTKNDYWMALAYNRSKLCN 300  
 QY 301 ILFSNELHRRILSPRGVTSNAVHPGNMYSNIHRSWVYTLFTLARPTKSM 352  
 DB 301 ILFSNELHRRILSPRGVTSNAVHPGNMYSNIHRSWVYTLFTLARPTKSM 352

RESULT 9

ABB63100

ID ABB63100 standard; protein; 409 AA.

XX ABB63100;

XX

DT 26-MAR-2002 (first entry)  
XX Drosophila melanogaster polypeptide SEQ ID NO 16092.  
DE  
XX Drosophila; developmental biology; cell signalling; insecticide;  
KW pharmaceutical.  
XX  
XX Drosophila melanogaster.  
OS  
XX WO200171042-A2.  
XX  
XX 27-SEP-2001.  
XX  
XX 23-MAR-2001; 2001WO-US009231.  
XX  
XX 23-MAR-2000; 2000US-0191637P.  
PR  
XX 11-JUL-2000; 2000US-00614150.  
XX  
XX (PEKE ) PE CORP NY.  
PA  
XX Venter JC, Adams M, Li PWD, Myers EW;  
PI  
XX WPI; 2001-656860/75.  
DR  
XX N-PSDB; ABL07203.  
XX  
XX New isolated nucleic acid detection reagent for detecting 1000 or more  
PT genes from Drosophila and for elucidating cell signalling and cell-cell  
PT interactions.  
XX  
XX Disclosure; SEQ ID NO 16092; 2lpp + Sequence Listing; English.  
PS  
XX  
XX The invention relates to an isolated nucleic acid detection reagent  
CC capable of detecting 1000 or more genes from Drosophila. The invention is  
CC useful in developmental biology and in elucidating cell signalling and  
CC cell-cell interactions in higher eukaryotes for the development of  
CC insecticides, therapeutics and pharmaceutical drugs. The invention  
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA  
CC sequences (ABL01840-ABL16175) and the encoded proteins (AB857737-  
CC AB872072). The sequence data for this patent did not form part of the  
CC printed specification, but was obtained in electronic format directly  
CC from WIPO at ftp.wipo.int/pub/published\_pct\_sequences  
XX  
XX Sequence 409 AA;  
XX  
Query Match 46.48; Score 1023; DB 4; Length 409;  
Best Local Similarity 49.58; Pred. No. 3.4e-95;  
Matches 200; Conservative 68; Mismatches 126; Indels 10; Gaps 6;  
XX  
9 LDDTDSDELPPGWEERTTKDGVVYANHTTEKTQWEHPKTKGKRKRVAGDLPYGEQETD 68  
DB 4 LPTDSEDELPPGWEERATDDGTVCYVNOQKTSQWTHPTGRSKRITGELPLGWEXYD 63  
XX  
69 ENGQVF-FVDHINKRTYLPRIAFVDDNPTK--PTTRQYDGGSTTAMEILQGRDPTGK 125  
DB 64 EQGKRFNFKETQORTNVDPRLAFAVEE-PTQNVAVQVRQFDSCTALQVLHGKDLHGR 122  
XX  
126 VVVVTGANSIGGETAKSFAHGHVILACRNARASEAVSRILEEWHKAKVEA--MTLD 183  
DB 123 TALITGANCIGYETARSLAHGCEIIFACNRSSAEAAIERIAQERPAARSRCAALD 182  
XX  
184 LALIRSVQHFAEAFKAKNVLPHVLVNAATFALPWSLTGDKLETTFOVNHGLHYLVQLL 243  
DB 183 LSSLSRVQFVEETIKQSVSHIDYILNAGVPAFYTRTVGLETTFQVSHLSHYLTQL 242  
XX  
244 QDVLCSAPARVIVVSSSEHRTDINDSLGKLDPSRLSPTKNDYMWAMLYNRSKLCNLF 303  
DB 243 ETLF--DYKTRIIVLSSSESRFANL--PVENLAVHLSPPPEKWSWYNNNAKLCNVLF 298  
XX  
304 SNELHRLSRPGVTSNAVHPCNMYSNIHRSWVYTLFLIARPTKSMOOGATTYCA 363  
DB 299 AQELAQKWKQSGVFSFLHPGNVSSLSRNYWYFRLLFIVRFTKSLQOAAATSYCA 358  
XX  
364 AVPELEGGLGMYFNCCRCMPSPAQGEETARTLWALSERLIQE 407

DB 359 TANELTGLSLGYNNCFCEPSKLSKSAALQQQLWKLSENLIAE 402  
XX  
RESULT 10  
AAB84486  
ID AAB84486 standard; protein; 189 AA.  
XX  
XX AAB84486;  
AC  
XX 05-SEP-2001 (first entry)  
DT  
XX DE Amino acid sequence of FRA16D oxidoreductase (FOR) III gene cDNA.  
XX  
XX Cancer associated protein; FOR gene; FRA16D; fragile site; aphidicolin;  
KW chromosomal rearrangement; cancer; splice variant; DNA instability;  
KW FRA16D oxidoreductase; neoplasia.  
XX  
XX Homo sapiens.  
OS  
XX WO200144466-A1.  
PN  
XX 21-JUN-2001.  
PD  
XX 15-DEC-2000; 2000WO-AUG01539.  
PF  
XX 16-DEC-1999; 99AU-00004711.  
PR  
XX 19-APR-2000; 2000AU-00007025.  
XX  
XX (WOME-) WOMEN'S & CHILDREN'S HOSPITAL.  
PA  
XX Richards R, Ried K, Finniss M, Hobson L, Mangeltsdorf M, Dayan S;  
PI Nancarrow J, Woollatt E, Baker E;  
PI  
XX WPI; 2001-398151/42.  
DR  
XX N-PSDB; AAH27868.  
XX  
PT Novel isolated 16q23.2 nucleic acid molecule, FRA16D oxidoreductase (FOR)  
PT gene associated with FRA16D site, useful for early diagnosis and  
PT assessment of risk of cancers associated with the FRA16D region.  
XX  
XX Claim 31; Fig 9; 150pp; English.  
PS  
XX  
XX The present sequence represents a human FRA16D oxidoreductase (FOR) III  
CC transcrit. The FOR gene encodes a cancer associated protein. The FRA16D  
CC site is a fragile site induced by aphidicolin, which is located within  
CC the FOR gene. The fragile site is the location of breakpoints of a  
CC variety of chromosomal rearrangements, and other mutations associated  
CC with cancers. The FOR protein is expressed as a number of splice  
CC variants. FOR gene polynucleotide fragments are capable of acting as  
CC specific primers or probes for detecting cancer associated variations of  
CC DNA sequence such as a point mutation or small DNA rearrangement  
CC associated with the tumour, a breakpoint of one or more chromosomal  
CC rearrangements associated with the tumour and a pause site within the  
CC FRA16 gene. FOR nucleic acid molecules are useful as markers to identify  
CC relationship between the fragile site (FRA16D) and the DNA instability in  
CC neoplasia which allows better diagnosis of cancers associated with the  
CC region  
XX  
SQ Sequence 189 AA;  
XX  
Query Match 42.1%; Score 927; DB 4; Length 189;  
Best Local Similarity 100.0%; Pred. No. 6.7e-86;  
Matches 172; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
XX  
1 MAALRYAGLDDTDSDELPPGWEERTTKDGVVYANHTTEKTQWEHPKTKGKRKRVAGDLP 60  
DB 1 MAALRYAGLDDTDSDELPPGWEERTTKDGVVYANHTTEKTQWEHPKTKGKRKRVAGDLP 60  
XX  
61 YGHEQETDENGQVFFVDHINKRTYLPRIAFVDDNPTKPTTRQYDGGSTTAMEILQGR 120  
DB 61 YGHEQETDENGQVFFVDHINKRTYLPRIAFVDDNPTKPTTRQYDGGSTTAMEILQGR 120

Qy 121 DFTGKVVVVVTGANGSIGFETAKSFALHGAHVILACRNMARASEAVSRILEEW 172  
Db 121 DFTGKVVVVVTGANGSIGFETAKSFALHGAHVILACRNMARASEAVSRILEEW 172

RESULT 11  
ABBI0434  
ID ABB10434 standard; protein; 173 AA.  
AC ABB10434;  
XX  
XX  
DT 10-JAN-2002 (first entry)  
DE Human cDNA SEQ ID NO: 742.  
XX  
XX Human; gene therapy; neural disorder; immune system disorder;  
KW muscular disorder; reproductive disorder; gastrointestinal disorder;  
KW pulmonary disorder; cardiovascular disorder; renal disorder;  
KW proliferative disorder; inflammation.  
XX  
OS Homo sapiens.  
XX WO200154474-A2.  
PN  
XX  
PD 02-AUG-2001.  
PF  
PF 17-JAN-2001; 2001WO-US001349.  
XX  
XX 31-JAN-2000; 2000US-0179065P.  
PR 04-FEB-2000; 2000US-0180628P.  
PR 24-FEB-2000; 2000US-0184664P.  
PR 02-MAR-2000; 2000US-0186350P.  
PR 16-MAR-2000; 2000US-0189874P.  
PR 17-MAR-2000; 2000US-0190076P.  
PR 18-APR-2000; 2000US-019123P.  
PR 19-MAY-2000; 2000US-0205515P.  
PR 07-JUN-2000; 2000US-0209467P.  
PR 28-JUN-2000; 2000US-0214886P.  
PR 30-JUN-2000; 2000US-0215135P.  
PR 07-JUL-2000; 2000US-0216647P.  
PR 07-JUL-2000; 2000US-021680P.  
PR 11-JUL-2000; 2000US-0217487P.  
PR 11-JUL-2000; 2000US-0217496P.  
PR 14-JUL-2000; 2000US-0218290P.  
PR 26-JUL-2000; 2000US-0220963P.  
PR 26-JUL-2000; 2000US-0220964P.  
PR 14-AUG-2000; 2000US-0224518P.  
PR 14-AUG-2000; 2000US-0224519P.  
PR 14-AUG-2000; 2000US-0225213P.  
PR 14-AUG-2000; 2000US-0225214P.  
PR 14-AUG-2000; 2000US-0225266P.  
PR 14-AUG-2000; 2000US-0225267P.  
PR 14-AUG-2000; 2000US-0225268P.  
PR 14-AUG-2000; 2000US-0225270P.  
PR 14-AUG-2000; 2000US-0225447P.  
PR 14-AUG-2000; 2000US-0225757P.  
PR 14-AUG-2000; 2000US-0225758P.  
PR 14-AUG-2000; 2000US-0225759P.  
PR 18-AUG-2000; 2000US-0226279P.  
PR 22-AUG-2000; 2000US-0226681P.  
PR 22-AUG-2000; 2000US-0226868P.  
PR 23-AUG-2000; 2000US-0227182P.  
PR 30-AUG-2000; 2000US-0227009P.  
PR 01-SEP-2000; 2000US-0228924P.  
PR 01-SEP-2000; 2000US-0228927P.  
PR 01-SEP-2000; 2000US-0229343P.  
PR 01-SEP-2000; 2000US-0229344P.  
PR 01-SEP-2000; 2000US-0229345P.  
PR 05-SEP-2000; 2000US-0229509P.  
PR 05-SEP-2000; 2000US-0229513P.  
PR 06-SEP-2000; 2000US-0230437P.  
PR 06-SEP-2000; 2000US-0230438P.  
PR 08-SEP-2000; 2000US-0231242P.

08-SEP-2000; 2000US-0231243P.  
08-SEP-2000; 2000US-0231244P.  
08-SEP-2000; 2000US-0231413P.  
08-SEP-2000; 2000US-0231414P.  
08-SEP-2000; 2000US-0232080P.  
08-SEP-2000; 2000US-0232081P.  
12-SEP-2000; 2000US-0231968P.  
14-SEP-2000; 2000US-0232397P.  
14-SEP-2000; 2000US-0232398P.  
14-SEP-2000; 2000US-0232399P.  
14-SEP-2000; 2000US-0232400P.  
14-SEP-2000; 2000US-0232401P.  
14-SEP-2000; 2000US-0233063P.  
14-SEP-2000; 2000US-0233064P.  
14-SEP-2000; 2000US-0233065P.  
21-SEP-2000; 2000US-0234223P.  
21-SEP-2000; 2000US-0234274P.  
25-SEP-2000; 2000US-0234997P.  
25-SEP-2000; 2000US-0234998P.  
26-SEP-2000; 2000US-0235484P.  
27-SEP-2000; 2000US-0235834P.  
27-SEP-2000; 2000US-0235836P.  
29-SEP-2000; 2000US-0236327P.  
29-SEP-2000; 2000US-0236367P.  
29-SEP-2000; 2000US-0236368P.  
29-SEP-2000; 2000US-0236369P.  
29-SEP-2000; 2000US-0236370P.  
02-OCT-2000; 2000US-0236802P.  
02-OCT-2000; 2000US-0237037P.  
02-OCT-2000; 2000US-0237038P.  
02-OCT-2000; 2000US-0237039P.  
02-OCT-2000; 2000US-0237040P.  
13-OCT-2000; 2000US-0239935P.  
13-OCT-2000; 2000US-0239937P.  
20-OCT-2000; 2000US-0240960P.  
20-OCT-2000; 2000US-0241221P.  
20-OCT-2000; 2000US-0241785P.  
20-OCT-2000; 2000US-0241786P.  
20-OCT-2000; 2000US-0241787P.  
20-OCT-2000; 2000US-0241808P.  
20-OCT-2000; 2000US-0241809P.  
20-OCT-2000; 2000US-0241825P.  
01-NOV-2000; 2000US-0244617P.  
08-NOV-2000; 2000US-0246474P.  
08-NOV-2000; 2000US-0246475P.  
08-NOV-2000; 2000US-0246476P.  
08-NOV-2000; 2000US-0246477P.  
08-NOV-2000; 2000US-0246478P.  
08-NOV-2000; 2000US-0246523P.  
08-NOV-2000; 2000US-0246524P.  
08-NOV-2000; 2000US-0246525P.  
08-NOV-2000; 2000US-0246526P.  
08-NOV-2000; 2000US-0246527P.  
08-NOV-2000; 2000US-0246528P.  
08-NOV-2000; 2000US-0246532P.  
08-NOV-2000; 2000US-0246609P.  
08-NOV-2000; 2000US-0246610P.  
08-NOV-2000; 2000US-0246611P.  
08-NOV-2000; 2000US-0246613P.  
17-NOV-2000; 2000US-0249207P.  
17-NOV-2000; 2000US-0249208P.  
17-NOV-2000; 2000US-0249209P.  
17-NOV-2000; 2000US-0249210P.  
17-NOV-2000; 2000US-0249211P.  
17-NOV-2000; 2000US-0249212P.  
17-NOV-2000; 2000US-0249213P.  
17-NOV-2000; 2000US-0249214P.  
17-NOV-2000; 2000US-0249215P.  
17-NOV-2000; 2000US-0249216P.  
17-NOV-2000; 2000US-0249217P.  
17-NOV-2000; 2000US-0249218P.  
17-NOV-2000; 2000US-0249244P.  
17-NOV-2000; 2000US-0249245P.





PR 14-SEP-2000; 2000US-0233065P.  
PR 21-SEP-2000; 2000US-0234223P.  
PR 21-SEP-2000; 2000US-0234274P.  
PR 25-SEP-2000; 2000US-0234997P.  
PR 25-SEP-2000; 2000US-0234998P.  
PR 26-SEP-2000; 2000US-0235484P.  
PR 27-SEP-2000; 2000US-0235834P.  
PR 27-SEP-2000; 2000US-0235836P.  
PR 29-SEP-2000; 2000US-0236327P.  
PR 29-SEP-2000; 2000US-0236367P.  
PR 29-SEP-2000; 2000US-0236368P.  
PR 29-SEP-2000; 2000US-0236369P.  
PR 29-SEP-2000; 2000US-0236370P.  
PR 02-OCT-2000; 2000US-0236803P.  
PR 02-OCT-2000; 2000US-0237037P.  
PR 02-OCT-2000; 2000US-0237038P.  
PR 02-OCT-2000; 2000US-0237039P.  
PR 02-OCT-2000; 2000US-0237040P.  
PR 13-OCT-2000; 2000US-0239935P.  
PR 13-OCT-2000; 2000US-0239937P.  
PR 20-OCT-2000; 2000US-0240960P.  
PR 20-OCT-2000; 2000US-0241221P.  
PR 20-OCT-2000; 2000US-0241783P.  
PR 20-OCT-2000; 2000US-0241786P.  
PR 20-OCT-2000; 2000US-0241787P.  
PR 20-OCT-2000; 2000US-0241808P.  
PR 20-OCT-2000; 2000US-0241809P.  
PR 01-NOV-2000; 2000US-0244617P.  
PR 08-NOV-2000; 2000US-0246474P.  
PR 08-NOV-2000; 2000US-0246475P.  
PR 08-NOV-2000; 2000US-0246476P.  
PR 08-NOV-2000; 2000US-0246477P.  
PR 08-NOV-2000; 2000US-0246478P.  
PR 08-NOV-2000; 2000US-0246523P.  
PR 08-NOV-2000; 2000US-0246524P.  
PR 08-NOV-2000; 2000US-0246525P.  
PR 08-NOV-2000; 2000US-0246526P.  
PR 08-NOV-2000; 2000US-0246527P.  
PR 08-NOV-2000; 2000US-0246528P.  
PR 08-NOV-2000; 2000US-0246532P.  
PR 08-NOV-2000; 2000US-0246609P.  
PR 08-NOV-2000; 2000US-0246610P.  
PR 08-NOV-2000; 2000US-0246611P.  
PR 08-NOV-2000; 2000US-0246613P.  
PR 17-NOV-2000; 2000US-0249207P.  
PR 17-NOV-2000; 2000US-0249208P.  
PR 17-NOV-2000; 2000US-0249209P.  
PR 17-NOV-2000; 2000US-0249210P.  
PR 17-NOV-2000; 2000US-0249211P.  
PR 17-NOV-2000; 2000US-0249212P.  
PR 17-NOV-2000; 2000US-0249213P.  
PR 17-NOV-2000; 2000US-0249214P.  
PR 17-NOV-2000; 2000US-0249215P.  
PR 17-NOV-2000; 2000US-0249216P.  
PR 17-NOV-2000; 2000US-0249217P.  
PR 17-NOV-2000; 2000US-0249218P.  
PR 17-NOV-2000; 2000US-0249244P.  
PR 17-NOV-2000; 2000US-0249245P.  
PR 17-NOV-2000; 2000US-0249264P.  
PR 17-NOV-2000; 2000US-0249265P.  
PR 17-NOV-2000; 2000US-0249297P.  
PR 17-NOV-2000; 2000US-0249299P.  
PR 17-NOV-2000; 2000US-0249300P.  
PR 01-DEC-2000; 2000US-0250160P.  
PR 01-DEC-2000; 2000US-0250391P.  
PR 05-DEC-2000; 2000US-0251030P.  
PR 05-DEC-2000; 2000US-0251988P.  
PR 05-DEC-2000; 2000US-0256719P.  
PR 06-DEC-2000; 2000US-0251479P.  
PR 08-DEC-2000; 2000US-0251856P.  
PR 08-DEC-2000; 2000US-0251868P.  
PR 08-DEC-2000; 2000US-0251869P.

PR 08-DEC-2000; 2000US-0251989P.  
PR 08-DEC-2000; 2000US-0251990P.  
PR 11-DEC-2000; 2000US-0254097P.  
PR 05-JAN-2001; 2001US-0259678P.  
XX (HUMA-) HUMAN GENOME SCI INC.  
XX Rosen CA, Barash SC, Ruben SM;  
XX MPI; 2001-451936/48.  
XX N-PSDB; AAS29665.  
XX Isolated polypeptide for treating, preventing and/or prognosing  
XX disorders of the endocrine system such as reproductive disorders,  
XX endocrine cancers and also for testing and detection e.g. diagnosis.  
XX Claim 11; SEQ ID NO 391; 604pp; English.  
XX Sequences AAU18282-AAU18507 represent endocrine polypeptides of the  
XX invention. Endocrine polypeptides and their associated polynucleotides  
XX are useful in the diagnosis, treatment and prevention of various types of  
XX disorders in e.g. humans, mice, rabbits, goats, horses, cats, dogs,  
XX chickens or sheep. A pathological condition can be determined by  
XX determining the presence or absence of a mutation in an endocrine  
XX polynucleotide. The treatable disorders include autoimmune diseases such  
XX as rheumatoid arthritis, hyperproliferative disorders such as neoplasms  
XX of the breast or liver, cardiovascular disorders such as cardiac arrest,  
XX cerebrovascular disorders such as cerebral ischaemia, nervous system  
XX disorders such as Alzheimer's disease, infections caused by bacteria,  
XX viruses and fungi, ocular disorders such as corneal infection, endocrine  
XX disorders such as Crohn's disease, renal disorders such as  
XX glomerulonephritis and respiratory disorders such as asthma. The  
XX polypeptides can also be used to aid wound healing, to prevent skin aging  
XX tissues and in chemotaxis. The polypeptides can also be used as a food  
XX additive or preservative to increase or decrease storage capabilities.  
XX Note: The sequence data for this patent did not form part of the printed  
XX specification, but was obtained in electronic format directly from WIPO  
XX at ftp.wipo.int/pub/published\_pct\_sequences  
Query Match 39.2%; Score 862.5; DB 4; Length 173;  
Best Local Similarity 92.5%; Pred. No. 2.3e-79;  
Matches 161; Conservative 2; Mismatches 10; Indels 1; Gaps 1;  
QY 3 ALRYAGLDLDTSEDLPGWEERTTKDQWVYANHTKTEKQWEHPTKGRKXVAGDLPYG 62  
Db 1 ALRYAGLDLDTSEDLPGWEERTTKDQWVYANHTKTEKQWEHPTKGRKXVAGDLPYG 60  
QY 63 WEQETDENGQVFFVDHINKRTTYLDPLAFTVDDNPTKPTTQRQYDGGSTTAMEILQGRDX 122  
Db 61 WEQETDENGQVFFVDHINKRTTYLDPLAFTVDDNPTKPTTQRQYDGGSTTAMEILQGRDX 120  
QY 123 TGVVVVVTGANGSIGGFETAKSFALHGAHVILACRNMARASEAVSRILEWHKAK 176  
Db 121 TGVVVVVTGANGSIGGFETAKSFALHGAHVILACRNMEXXNNVT-ILEWHKXKR 173  
RESULT 13  
ABP67021  
ID ABP67021 standard; protein; 173 AA.  
XX  
AC ABP67021;  
XX  
DT 09-DEC-2002 (first entry)  
XX  
DE Human polypeptide SEQ ID NO 742.  
XX  
KW Human; nootropic; neuroprotective; cytostatic; dermatological; virucide;  
KW immunosuppressive; antiinflammatory; anti-HIV; antibacterial; vulnery;  
KW antiparkinsonian; antiscikling; antianaemic; antiarthritic; cancer;  
KW antirheumatic; hepatotropic; cerebroprotective; antiinflammatory;  
KW antiallergic; antidiabetic; antitumor; anticonvulsant; antifungal;

antiparasitic; cardiant; immune disorder; cardiovascular disorder;  
neurological disease; infection; nephrotropic; gene therapy; vaccine.

Homo sapiens.

US2002090672-A1.

11-JUL-2002.

17-JAN-2001; 2001US-00764853.

31-JAN-2000; 2000US-0179065P.

04-FEB-2000; 2000US-0180628P.

28-JUN-2000; 2000US-0214886P.

07-JUL-2000; 2000US-0216647P.

07-JUL-2000; 2000US-0216880P.

11-JUL-2000; 2000US-0217487P.

11-JUL-2000; 2000US-0217496P.

14-JUL-2000; 2000US-0218290P.

26-JUL-2000; 2000US-0220963P.

26-JUL-2000; 2000US-0220964P.

14-AUG-2000; 2000US-0224518P.

14-AUG-2000; 2000US-0224519P.

14-AUG-2000; 2000US-0225267P.

14-AUG-2000; 2000US-0225268P.

14-AUG-2000; 2000US-0225270P.

14-AUG-2000; 2000US-0225447P.

14-AUG-2000; 2000US-0225757P.

14-AUG-2000; 2000US-0225758P.

22-AUG-2000; 2000US-0226688P.

30-AUG-2000; 2000US-0228924P.

01-SEP-2000; 2000US-0229287P.

01-SEP-2000; 2000US-0229343P.

01-SEP-2000; 2000US-0229344P.

01-SEP-2000; 2000US-0229345P.

05-SEP-2000; 2000US-0229509P.

08-SEP-2000; 2000US-0229513P.

21-SEP-2000; 2000US-0231413P.

21-SEP-2000; 2000US-0234223P.

21-SEP-2000; 2000US-0234274P.

25-SEP-2000; 2000US-0234997P.

27-SEP-2000; 2000US-0235834P.

29-SEP-2000; 2000US-0236372P.

29-SEP-2000; 2000US-0236367P.

29-SEP-2000; 2000US-0236368P.

29-SEP-2000; 2000US-0236369P.

23-SEP-2000; 2000US-0236370P.

02-OCT-2000; 2000US-0236802P.

02-OCT-2000; 2000US-0237037P.

02-OCT-2000; 2000US-0237038P.

02-OCT-2000; 2000US-0237039P.

13-OCT-2000; 2000US-0239935P.

20-OCT-2000; 2000US-0240960P.

20-OCT-2000; 2000US-0241785P.

20-OCT-2000; 2000US-0241809P.

01-NOV-2000; 2000US-0244617P.

17-NOV-2000; 2000US-0249239P.

08-DEC-2000; 2000US-0251856P.

08-DEC-2000; 2000US-0251868P.

08-DEC-2000; 2000US-0251869P.

(ROSE/) ROSEN C A.

(RUBE/) RUBEN S M.

(BARA/) BARASH S C.

Rosen CA, Ruben SM, Barash SC;

WPI; 2002-681727/73.

N-PSDB; ABV83993.

Novel polypeptide useful for diagnosis, prognosis, prevention, and  
treatment of immune, hyperproliferative, renal, respiratory,

cardiovascular, reproductive, endocrine, gastrointestinal and  
neurological disorders.

Claim 11; SEQ ID NO 742; 369pp + Sequence Listing; English.

The invention relates to novel genes (ABV83682-ABV84101) and proteins  
(ABP66710-ABP67129) useful for preventing, treating or ameliorating  
medical conditions e.g. by protein or gene therapy. The genes are  
isolated from a range of human tissues disclosed in the specification.  
The nucleic acids, proteins, antibodies and (ant)agonists are useful in  
the diagnosis, treatment and prevention of: (a) cancer, e.g. breast and  
ovarian cancer and other cancers of the adrenal gland, bone, bone marrow,  
breast, gastrointestinal tract, liver, lung, or urogenital; (b) immune  
disorders e.g. Addison's disease, allergies, autoimmune haemolytic  
anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's disease,  
multiple sclerosis, rheumatoid arthritis and ulcerative colitis; (c)  
cardiovascular disorders such as myocardial ischaemia; (d) wound healing  
; (e) neurological diseases e.g. cerebral anoxia and epilepsy; and (f)  
infectious diseases such as viral, bacterial, fungal and parasitic  
infections. Note: The sequence data for this patent did not form part of  
the printed specification, but was obtained in electronic format directly  
from WIPO at ftp.wipo.int/pub/published\_pct\_sequences

Sequence 173 AA;

Query Match 39.2%; Score 862.5; DB 5; Length 173;

Best Local Similarity 92.5%; Pred. No. 2.3e-79;

Matches 161; Conservative 2; Mismatches 10; Indels 1; Gaps 1;

QY 3 ALRYAGLDDTSDSELPGWERTTKDGVVYVYANHTTEKTQWEHPKTKRKRVAGDLPYG 62

Db 1 ALRYAGLDDTSDSELPGWERTTKDGVVYVYANHTTEKTQWEHPKTKRKRVAGDLPYG 60

QY 63 WEQETDENGQVFFVDHINKETTYLDPLAFTVDDNPKPTTRQYDGSSTAMEILQGRDF 122

Db 61 WEQETDENGQVFFVDHINKETTYLDPLAFTVDDNPKPTTRQYDGSSTAMEILQGRDX 120

QY 123 TGKVVVVTGANSIGIGFETAKSFALHGAHVILACNMAVSEAVSRILEENHAK 176

Db 121 TGKVVVVTGANSIGIGFETAKSFALHGAHVILACNMEVXPNNVT-ILEEWXXR 173

RESULT 14

AAE14438

ID AAE14438 standard; protein; 316 AA.

XX AAE14438;

AC AAE14438;

XX 26-MAR-2002 (first entry)

DE Human drug metabolising enzyme (DME)-1.

XX Human; drug metabolising enzyme; DME-1; autoimmune; inflammatory;  
cell proliferative; developmental; endocrine; eye; metabolic; AIDS;  
gastrointestinal disorder; liver disorder; cancer; arteriosclerosis;  
adult respiratory distress syndrome; anaemia; epilepsy; hypothyroidism;  
hypochalams; pituitary; diabetes; hypogonadism; conjunctivitis;  
glaucoma; cystic fibrosis; hypercholesterolaemia; gastritis;  
peptic ulcer; hepatitis; gene therapy.

OS Homo sapiens.

XX Key Location/Qualifiers

FT Domain 1..19

XX /label= Transmembrane\_domain

XX WO200190334-A2.

XX 29-NOV-2001.

PD 25-MAY-2001; 2001WO-US017150.

XX 25-MAY-2000; 2000US-0207901P.

QY	281	SFTKNDY	WAMLAYNR	SKCNIL	FNELHRR	LSRPGV	TSNAVH	PGNMWY	GNIHRS	WWVYTL	340	
Db	189	QSEKR	-YSRGPAY	CHSKLANV	LFRELAKR	LQGTGVT	TYAVHFG	-VVRSEL	VHRSHS	ILICL	246	
QY	341	LFTLAR	PETKSMQO	GAATTV	CAAVPE	LGGMY	FNNCCR	CPMSP	EAQSE	EARTL	400	
Db	247	LWRLSP	VPVKARE	GAQTS	LHCAL	AEGLP	SGKIF	SDCKR	TWVSP	ARNKTA	ERLWN	306
QY	401	SERLI	405									
Db	307	SCELF	311									

ADAS4192  
ID ADAS4192 standard; protein; 316 AA.  
XX  
AC ADAS4192;

DE Human protein, SEQ ID 1760.  
XX  
XX  
KW Cytostatic; Anti-inflammatory; Osteopathic; Neuroprotective; Nootropic;  
DE Gene Therapy; human; secretory protein; membrane proteins; cancer;  
KW inflammatory disease; osteoporosis; neurological disease.  
XX

CC (ADA54071). The coding regions are useful in the gene therapy of  
CC diseases caused by abnormalities of the proteins, e.g. cancer,  
XX inflammatory diseases, osteoporosis or neurological disease.  
XX  
SO Sequence 316 AA:

Query Match 24.9%; Score 549.5; DB 6; Length 316;  
Best Local Similarity 41.6%; Pred. No. 5.3e-47;  
Matches 127; Conservative 49; Mismatches 116; Indels 13; Gaps 4

QY 101 PTTQRQYDGSITTAEMILQLGRDFTGKVVVVGANSIGFETAKSFALHGAFVILACRNMAR 160  
:  
Db 20 PSIRKFAGGVCRINV---QLPGKVVTIGANTGICKETARELASRGARVVIACEDVLK 75  
:  
QY 161 ASEAVSRILEEHWKHAKEAMTLDIALRSVOHFPAEFKKNVPELHVLVCNAATFALPWSL 220  
: :

Db	76	GESAASEIRVDTKNSQVLVRKLDLSDTKSIRAPAEGLAEKQLHILINNAGVMMCPYSK	135
QY	221	TXDGLTTFQVNHGFIYVQLQDVLCSAPARVIVSSSEHRTDINDSLGKLDSESL	280
Db	136	TADGFETHLGVNHLGHFLTYLLLEKVKVSAPARVVNSVAHH-----IGKIFHDL	188
QY	281	SPTKNDYMWAMLAYNRSKLCNILLFSNELHRLSPRGVTSNAVHPGNMNYSNIHRSWWVYTL	340
Db	189	QSEKR-YSRGFAYCHSKLANVLFTRELAKRLQGTGVTYAVHFG-VVRSELVRHSSLLCL	246
QY	341	LFTLARPFTKSMQOGAATTYYCAVPELGLGGMFNCCRCMPSPSAOSEETARTLWAL	400
Db	247	LWLFSPFVKTAREGAQTSIHCAEGLPLESGKYFSDCKRTWVSPARNNKTAERLWNV	306
QY	401	SEELI 405	
Db	307	CELL 311	

Search completed: May 24, 2004, 10:00:09  
Job time : 61 secs